CAN COMPLEMENTARY AND ALTERNATIVE MEDICINE BE EVALUATED WITHIN THE FRAMEWORK OF EVIDENCE-BASED MEDICINE?

by

Jeremy S. Guttman

B.A., University of Chicago, 2005

Submitted to the Graduate Faculty of

The Kenneth P. Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Master of Arts in Bioethics

University of Pittsburgh

2014

UNIVERSITY OF PITTSBURGH

THE KENNETH P. DIETRICH SCHOOL OF ARTS & SCIENCES

This thesis was presented

by

Jeremy S. Guttman

It was defended on

November 18, 2014

and approved by

Gary Fischer, MD, Associate Professor of Medicine, Division of General Internal Medicine

Ronald Glick, MD, Assistant Professor, Psychiatry, Physical Medicine and Rehabilitation, University of Pittsburgh School of Medicine

Kenneth Schaffner, MD, PhD, University Professor, Department of History and Philosophy of Science

Thesis Advisor: Kenneth Schaffner, MD, PhD, University Professor, Department of History and Philosophy of Science

CAN COMPLEMENTARY AND ALTERNATIVE MEDICINE BE EVALUATED WITHIN THE FRAMEWORK OF EVIDENCE-BASED MEDICINE?

Jeremy S. Guttman, M.A.

University of Pittsburgh, 2014

There are many complications in fitting complementary and alternative medicine (CAM) into the hierarchy of evidence-based medicine, which favors "gold standard" RCTs and meta-analyses. Some argue that CAM practices that have been practiced for many years don't need rigorous studies to verify them, or that the most rigorous scientific studies aren't necessary as long as people get better. However, success over many years is not enough to validate a treatment. Also, pragmatic outcomes-focused arguments sidestep philosophical and scientific issues about whether the therapies "do anything" at all beyond non-specific and placebo effects. There are a number of reasons why randomization and placebo controls may not be appropriate to accurately study CAM therapies, such as studies where a subjects' effects are dampened by the act of participating in a study, studies where a sham control is more effective than the best conventional treatment, or studies that miss important subsets. Some argue that CAM should meet the same high standards as conventional medicine. However, 1. This holds CAM to a double standard because most conventional medical practices have less-than-rigorous evidence to support them. 2. The gold standard inherent in the evidence-based medicine hierarchy is not always appropriate. 3. CAM resembles a scientific paradigm and it may not be possible for scientists in the current biomedical paradigm to understand and evaluate CAM. And 4. Even if CAM can be accommodated within the evidence-based medicine framework, the attitude and priorities of the evidence-based medicine movement are problematic. Less rigorous studies are not often

considered to be acceptable types of evidence, and there are few financial incentives to study CAM treatments in a more focused manner.

TABLE OF CONTENTS

1.0 INTRODUCTION

1.1 EBM OVERVIEW

Evidence-based medicine (EBM) is a movement that has gained traction in the medical field since the 1990s. According to an important 1996 article on EBM, "Evidence-based Medicine is the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients" (Sackett 1996). EBM favors the reliance on common criteria and information from published studies of clinical trials to inform medical practice.

The purpose of the EBM movement is to buttress the scientific nature of medicine and reduce potential biases that would otherwise lead to improper care. EBM studies are supposed to give the practice of medicine a stronger evidence base with the hopes of reducing how much clinical medicine relies on heuristics and the unsystematic personal training and experience of physicians. EBM started out as an antidote to the inadequacies and unreliability of experience-based clinical expertise, but it was later softened to acknowledge the balance of evidence and expertise needed in real-life clinical practice. The original definition spoke of "[de-emphasizing] intuition, unsystematic clinical experience, and pathophysiological rationale" for clinical decision making (Guyatt et al 1992). Later, the founders of the EBM movement acknowledged that "[e]xternal clinical evidence can inform, but can never replace, individual clinical expertise" (Sackett et al, 1996). However, some EBM proponents generally perceive the practice of the "art

of medicine" and the use of clinical expertise to be inferior to the practice of evidence-based medicine. They think that medicine should be based on empirical evidence as much as possible because it is the best method to reduce bias. "In everyday speech evidence based medicine…is considered separate from or even antithetical to reliance on clinical expertise or patient values" (Henry et al 2007).

Many forms of EBM include some kind of hierarchy of evidence that puts pooling the results of different studies together at the top of the hierarchy. The hierarchies also generally favor the use of more rigorous studies that have more ways of reducing potential bias. Different versions of an EBM hierarchy have been presented by different organizations and working groups since the movement started. In 2002 a document prepared for the Agency for Healthcare Research and Quality reported that there were 40 different systems of rating in use (West et al 2002, Solomon 2011). Some believe the large variety of EBM hierarchies cause there to be too many possibilities for what constitutes evidence-based practice (Upshur 2003), but the typical evidence-based medicine hierarchy is as follows:

> Meta-Analyses and Systematic Reviews of Randomized, Controlled Trials (RCTs)
> Randomized, Controlled Trials (The so-called "Gold Standard")
> Other controlled clinical trials
> Observational studies, such as case-control and cohort studies
> Case studies, anecdote, bench studies and personal opinion
> (Greenhalgh 2010)

Randomized trials are trials that allocate trial subjects into each of the arms of a trial by random chance. Controlled trials compare the groups of subjects receiving an experimental treatment with a control arm such as a no-treatment or wait-list control group, an active control group that receives an active treatment such as usual care or standardized care, or a placebo control group that receives a treatment that resembles the experimental treatment but is inert and has no specific therapeutic effects beyond non-specific and placebo effects. One major trend in the

medical science community is that studies are not considered legitimate unless they meet the "gold standard" RCT criteria on the first or second tiers of the hierarchy. "Any study using the RCT design is considered superior to any study not using this design—and nowhere is this demonstrated more clearly than in EBM's method of evidence evaluation, whose evidence hierarchies feature RCTs above all other forms of evidence. As a consequence of such beliefs and practices, one has much more hope of getting a bad RCT funded than any caliber of, say, observational study, both because of the opinions of funding bodies and, less directly, because of the publication policies of journals. So there are clear incentives for researchers to conduct nothing but RCTs if they can possibly help it" (Grossman and Mackenzie 2005). It is not just that RCTs are considered to be superior to other types of studies; many people infer from the concept of "gold standard" evidence that RCTs are the only type of real scientific evidence in medicine (Worrall 2007) because they have the most safeguards against bias.

There have been a number of academic critiques of the EBM hierarchy on philosophical, statistical and practical grounds (Howick 2011, Worrall 2007, Bluhm 2009, Borgerson 2005).

> Critics argue that evidence produced by meta-analyses of RCTs is not particularly helpful for guiding treatment decisions at the level of individual patients (RCTs were, after all, originally designed for use in agriculture), that evidence supplied by qualitative research or case-studies can be, at times, even more helpful and relevant than the 'gold standard' evidence, that the direction of health care research is biased toward pharmaceuticals because the RCT methodology is most appropriate for evaluating pills, and that it does not make sense to have one evidence hierarchy, or one 'best methodology' in a field as diverse as medicine (Borgerson 2009).

Some critics have proposed modifications and more complex hierarchies to accommodate these critiques (Howick et al 2011). These alternative proposals and improvements have received some attention in the medical community and the US government has recently allocated significant funding for somewhat less rigorous comparative effectiveness research. However, the perception remains that RCTs are the gold standard above all else.

1.2 CAM OVERVIEW

Complementary and Alternative Medicine (CAM) encompasses a set of healing practices such as acupuncture and Traditional Chinese Medicine, chiropractic care, meditation, yoga, homeopathic medicine (vials of diluted substances), naturopathic medicine (focusing on foods, diet, herbs and nutritional supplements to heal the body), Southeast Asian Ayurvedic medicine, bodywork therapies (massage, craniosacral therapy), biofield healing practices (Reiki, Therapeutic Touch, Qigong) and indigenous folk medicine practices that are not commonly taught as part of conventional medical curriculums. Most CAM practices are based on the interconnectedness of the whole body (and the connected relationship between the mind and body) rather than honing in on specific diseases or specific body parts or body systems. Many CAM practices such as homeopathy, ayurvedic medicine and Traditional Chinese Medicine have elaborate theoretical bases that are not based in modern science and run counter to what is understood about the human body (and the universe) from orthodox biology, chemistry, and physics. (Hughes 2008). Many CAM modalities have multi-year educational programs where CAM practitioners are credentialed.

There has been a long history of tension between CAM and conventional medicine. The traditional medical community's past attempts to define CAM have tried to categorize the practices as those lacking scientific evidence or those not used by biomedical practitioners, but these definitions have "failed to provide a lasting definition of CAM" (Georgiou 2006). The

American Medical Association's code of ethics used to contain a "consultation clause" that stipulated that it was unethical to fraternize with or make referrals to "cult" practitioners which included various alternative practitioners (Kaptchuk and Miller 2005). In 1983 a court even ruled that the American Medical Association had engaged in a "nationwide conspiracy" to eliminate the chiropractic profession (Cohen 2004). Nowadays there is a somewhat more fluid relationship between CAM and conventional medicine. Many conventional practitioners will refer patients to CAM practitioners (Snyder 2007), oftentimes when all other available options have been exhausted.

Many forms of CAM are now referred to as Integrative Medicine to signify how patients often use both conventional medicine and CAM therapies at the same time, and how the two types of medicine need not be mutually exclusive. Two 2007 surveys conducted by the Centers for Disease Control and Prevention's National Center for Health Statistics found that 38% of the US population uses at least one form of CAM treatment at an annual cost of 33.9 billion American dollars (Barnes et al 2008, Nahin et al 2009).

1.3 ISSUES WITH EBM AND CAM

Most CAM practices have not been studied with many rigorous double-blinded placebo-controlled clinical studies. Even when they are studied, the evidence is often mixed. Since the mechanisms behind many CAM therapies are largely invisible and scientifically unproven, and since the theories behind them are not grounded in conventional science, one argument for why CAM therapies need to be studied within the framework of evidence-based medicine is to determine whether they "do anything" or "work" at all. Many people are skeptical about

whether CAM therapies function the way CAM practitioners claim, and these skeptics tend to suspect that CAM therapies are, or could be, elaborate placebos that are merely dressed up in rituals with unsound mechanistic explanations about Qi, acupuncture meridians, human biofields, chakras, water memory, vital energy, etc. Placebo-controlled trials are the commonly accepted way to distinguish if therapies "work" because the goal of a placebo-controlled trial is to distinguish the "specific" healing effects of the therapies themselves from any "non-specific" healing effects such as the placebo effect. Placebo-controlled trials are considered a rigorous way of testing for efficacy, i.e. if the treatment "does anything" at all. Placebos are especially useful tools in a clinical trial if the beneficial effects of a therapy are difficult to distinguish from placebo effects due to the nature of the therapy.

Testing for efficacy is different than testing for effectiveness and this distinction is important in debates about medical evidence. Effectiveness trials tend to be less rigid and controlled than efficacy trials and allow more leeway for the way a treatment is conducted. Effectiveness trials also include a wider variety of subjects who may have other medical conditions (comorbidities) or who may be getting other treatments at the same time. The purpose of effectiveness trials is to test for outcomes and how treatments would work in real-world settings, "for whom and under what conditions" while efficacy trials test for safety and efficacy, "what can work" (AcademyHealth 2009). In the United States, the Food and Drug Administration (FDA) requires most pharmaceutical drugs to undergo randomized, double-blind placebo-controlled testing to determine whether there is "substantial evidence" to support their effectiveness before they are approved by the FDA for use (Applications for FDA Approval to Market a New Drug, 2014). Unfortunately, it is much easier to conduct placebo-controlled trials

for pharmaceutical drugs than for CAM therapies. It is debatable whether it is necessary or even possible to conduct placebo-controlled trials for CAM therapies.

1.4 NON-SPECIFIC EFFECTS

Two terms used in this paper are specific effects and non-specific effects. Non-specific effects refer to the effects of a therapy that are happening at the same time or after a therapy but are independent from any specific therapeutic property of the therapy itself. These non-specific healing effects can be attributable to placebo effects (patient's response to the fact that s/he is being given a treatment), hawthorne effects (patient's response to being observed and assessed in a study), effects of the therapeutic setting, natural history effects (time, regression to the mean, spontaneous remission) and the effects of the expectations and interactions of the practitioner and patient (Kaptchuk et al 2008, Bouchet et al 1996, Lohr et al 2003). For example, if a patient had a shoulder pain problem and a physician gave the patient a Tic-Tac breath mint and told the patient that it was a shoulder pain pill, and then the patient's shoulder pain went away, the resolution and healing of the shoulder pain would likely be due to non-specific healing effects of time, attention, or expectations of healing rather than any specific biological, physical or chemical shoulder-pain-healing properties of the Tic-Tac breath mint itself.

Specific effects, on the other hand, refer to the therapeutic effects, if any, attributable to the actual therapy itself, above and beyond and independent from any non-specific and placebo effects. If someone has scurvy, and s/he takes vitamin C and then no longer has scurvy, the benefit was likely due to the specific therapeutic effects of the properties of the vitamin C rather than any placebo effects or other non-specific effects.

2.0 ARGUMENTS SOME MAKE FOR WHY CAM IS INCOMPATIBLE WITH THE FRAMEWORK OF EVIDENCE-BASED MEDICINE

2.1 THREE ARGUMENTS

There are a number of paradoxes and controversies related to applying the framework of evidence-based medicine to complementary and alternative medicine practices. Many unconventional medicine proponents and practitioners have tended to be wary of attempts to fit CAM modalities into a biomedical research framework (Barry 2006) and there is no consensus on the right way of assessing the efficacy of CAM (Hufford 2002). Three arguments are sometimes made to defend CAM from modern scientific evidence-gathering practices:

2.1.1 Tradition

One argument for the incompatibility of the EBM framework and CAM is an appeal to tradition, that if something like acupuncture or ayurvedic medicine has been practiced for thousands of years with success and repetition then it must have some legitimacy and rigorous trials are not necessary to prove it.

2.1.2 Pragmatism

Another argument for the incompatibility of the EBM framework and CAM is an appeal to pragmatism, that if patients consistently get better from a healing practice, it doesn't really matter if the therapy hasn't been conclusively shown to work beyond a placebo in rigorous, controlled scientific studies. In this pragmatic line of thinking, the mechanism of how people get better doesn't matter; the only thing that matters is the outcome: people tend to get better with a particular therapy on a consistent basis.

2.1.3 Incompatible with RCTs

A third argument is that double-blinded, randomized, controlled (often placebo-controlled) clinical trials of CAM therapies are either not possible or, when attempted, fail to find and validate effects that actually exist because the nature of CAM therapies combined with the nature of scientific testing make the effects of CAM therapies undiscoverable in the context of a rigorous clinical trial. The first two arguments are about why the framework of EBM should not be applied to CAM; this third argument is that framework of EBM cannot be applied CAM.

2.2 CRITICAL ANALYSIS OF FIRST TWO ARGUMENTS THAT DEFEND CAM FROM EBM

2.2.1 Tradition: If it's worked for thousands of years, it must be legitimate

One substantial counterexample to the appeal to tradition is Ptolemaic astronomy. Ptolemaic astronomy was based on the idea that the earth, rather than the sun, is the center of the universe. It was practiced for over 1500 years with success and repetition, but its entire premise turned out to be illegitimate once it was discovered that the earth was not the center of the universe. The theory behind most CAM therapies is not well-established enough to preclude the possibility that they could turn out to be faulty in this same way, despite how long they have been practiced with success. By itself, a long history of successful usage is not a valid argument to support a healing practice's legitimacy. In fact, appeal to tradition is considered one of the logical fallacies in elementary philosophy.

This appeal to tradition is often applied to acupuncture theory and Traditional Chinese Medicine because a particular acupuncture textbook dates back thousands of years. However, it turns out that the current systematization of Traditional Chinese Medicine actually only dates back to the 1950s, when Emporer Mao Zedong wanted to show off the gems of Chinese culture to the world. The idea that TCM is thousands of years old

> is a newly written history, constructed since the 1970s to provide a sense of noble lineage and unbroken continuity to an entity that had never clearly been a single 'thing' before…but instead had been many different practices and schools identified with everything from Taoist health and longevity practices to Buddhist meditative practices to martial arts. In the modern era, a certain kind of whole had been cobbled together out of varied selected parts. To be fitted into the whole, each of these parts had first to be lifted out of the local context in which it had originally developed…It then had to be dusted off and reworked in ways appropriate for the goals of modern China (Harrington 2001).

It is true that the practice of acupuncture and the basic tenets of acupuncture meridian theory date back thousands of years, but the specific way acupuncture is currently practiced and the way it is combined and associated with other Traditional Chinese Medicine practices is not so old.

2.2.2 Pragmatism: All that matters is that people heal and get better

The appeal to pragmatism throws to the side issues of mechanism and specific efficacy. It is often adopted by CAM practitioners as well as conventional practitioners who recommend CAM therapies. Practitioners make this argument in order to justify their decisions in light of a "lack of evidence".

This perspective essentially sidesteps a lot of philosophical and scientific issues that arise in controversies about CAM evidence and mechanism. For example, if CAM therapies don't really "do" anything on their own, should modern scientific medicine embrace and recommend therapies that are essentially elaborate and largely effective façades for the placebo effect? And if CAM therapies do work, why do rigorous attempts to study them scientifically so often fail to find the specific effects that do exist?

Sometimes this pragmatic attitude towards a "lack of evidence" in support of CAM practices is framed as an argument for why CAM isn't compatible with rigorous western science and evidence-based medicine (Walach 2003; Tonelli and Callahan 2001). However, it turns out that the conflict here is not as much about "lack of evidence" and complete scientific incompatibility as it is about a lack of highest standards scientific evidence that includes randomization, controls (including placebo controls), strict inclusion criteria and single or double-blindedness to maximally reduce potential biases and confounding factors. There are actually tens of thousands of clinical trials that have been performed on many CAM methods,

particularly for acupuncture (Ernst et al 2011). The problem is that the less rigorous trials tend to highlight the positive effects of the therapy, while the most rigorous placebo-controlled efficacy trials are often (but not always) null/inconclusive.

Pragmatism can actually be studied without the use of placebos using research methods that focus on effectiveness and real-world outcomes rather than efficacy in optimal conditions. Many supporters of CAM therapies have pushed for prioritizing pragmatic trials, including the National Center for Complementary and Alternative Medicine (NCCAM), part of the National Institute of Health (Briggs 2014). The problem is that these types of studies are sometimes not considered to be rigorous enough compared to placebo-controlled studies. So the real question to be explored here might not be whether the framework EBM can be applied to CAM at all, but whether the EBM hierarchy is appropriate for the nature of CAM therapies, and how important it is for there to be placebo-controlled studies conclusively proving that CAM therapies "do anything" beyond non-specific and placebo effects.

2.2.2.1 Pragmatic Comparative Effectiveness Trials

There has been a recent focus on effectiveness trials in United States health policy since 2009. The United States government earmarked $1 billion to study the comparative effectiveness of medical treatments as part of the 2009 American Recovery and Reinvestment Act "Stimulus Package". The Institute of Medicine defined Comparative Effectiveness Research (CER) as "the generation and synthesis of evidence that compares the benefits and harms of alternative methods to prevent, diagnose, treat, and monitor a clinical condition, or to improve the delivery of care. The purpose of CER is to assist consumers, clinicians, purchasers, and policy makers to make informed decisions that will improve health care at both the individual and population

levels" (Ratner et al 2009). An independent nongovernmental nonprofit organization called the Patient-Centered Outcomes Research Initiative (PCORI) was also established as part of the 2010 "Obamacare" Patient Protection and Affordable Care Act in 2010. It will fund billions of dollars in outcomes research for medical treatments as well as other healthcare services through 2019. The types of outcomes research promoted by PCORI overlaps with many components of comparative effectiveness research (CER), but PCORI's funding criteria is somewhat broader and includes additional focuses and research methodologies (Patient-Centered Outcomes Research Definition Revision: Response to Public Input, 2012).

Pragmatic research trials are one experimental way of conducting comparative effectiveness research. Different treatments are usually compared against each other to test for effectiveness in real world conditions in order to enhance external validity and generalizability. "Such studies use wide inclusion criteria, allow flexibility in treatment regimen and often focus on patient-centered outcomes such as quality of life or self-recorded pain" (Lewith et al 2010). Even though they have wide exclusion criteria, some pragmatic trials can still be placed at the top of the evidence-based medicine hierarchy as pragmatic RCTs that compare the outcomes of a treatment to the outcomes of control treatments such as standard/usual care (which would be considered an "active control" treatment instead of a placebo control or a no-treatment control). Other pragmatic trials compare the outcomes of different treatments to each other without having a control arm.

An example of a pragmatic RCT is a 2013 PCORI-funded trial to study the effects of glucose monitoring on people who have diabetes but do not use insulin to treat it. The trial is underway, to be completed in July 2016, and planning to randomize 450 subjects to one of three arms: 1. self-monitoring of blood glucose with standard feedback of glucose values from a

glucose meter, 2. self-monitoring of blood glucose with enhanced feedback including glucose values as well as automated, tailored feedback messaging, and 3. the control arm, education about maintaining glucose levels without any self-monitoring of glucose levels with a glucose meter (Three Approaches to Glucose Monitoring in Non-insulin Treated Diabetes 2014). Another example of a pragmatic comparative effectiveness RCT is the PROspective Multicenter Imaging Study for Evaluation of Chest Pain (PROMISE) trial which tests 1. anatomical testing with coronary CT angiography against 2. a usual care control arm of functional stress tests for subjects presenting with symptoms of coronary artery disease (PROspective Multicenter Imaging Study for Evaluation of Chest Pain (PROMISE) 2014).

In addition to more rigorous RCTs, there are many non-experimental ways of conducting comparative effectiveness research, such as retrospective cohort studies, a type of non-randomized observational study that looks backwards at medical records or insurance records. Sometimes a process known as matching can be employed in observational trials to partially simulate the way randomization reduces biases in clinical trials. Matching tries to balance each group in an observational study with respect to possible relevant factors such as age, race and gender so that, for example, the mean age in a treated and control group will be similar. Comparative effectiveness studies are sometimes perceived to primarily consist of prospective, forward-looking randomized clinical trials that compare treatments against each other. However, because retrospective backwards-looking observational studies are easier and less expensive to conduct, and because observational studies still fall under the umbrella of comparative effectiveness research, the majority of comparative effectiveness studies are not randomized (Comparative Effectiveness Research 2014).

3.0 INCOMPATIBILITY WITH RCTS: FAILING TO FIND THE EFFECTS OF CAM THAT DO EXIST

I am now going to discuss some of the problems that arise in applying highest standards scientific evidence to CAM in the form of randomized, controlled (often placebo controlled) trials. In my discussion of these issues I am going to use acupuncture as a primary example. The first part of this section discusses paradigm issues and problems with randomization, the second part discusses the problems with placebo controls, and the third part discusses subset problems.

3.1 THOMAS KUHN AND PROBLEMS WITH RANDOMIZATION AND BELIEFS

The philosophical and epistemological issues that arise with the framework of EBM and CAM share many similarities with Thomas Kuhn's history and philosophy of science theory about scientific paradigms. Kuhn's theory distinguishes scientific paradigms that have different quasi-metaphysical commitments, different types of data and different puzzles to be solved. Kuhn documents how scientific anomalies that do not fit into a prevailing paradigm will slowly build up over time, but "normal scientists" within a paradigm will not worry about the anomalies until a new pre-paradigm is formed to better account for all the anomalies. Kuhn discusses how scientists who are beholden an older paradigm will resist the new one. Eventually there is a

scientific revolution and the new paradigm supersedes the old one. However, many of the scientists who are beholden to the older paradigm will never come around to accept the new paradigm. Some historical examples of theories that upset older pre-existing paradigms and led to new paradigms getting adopted include Copernican astronomy, Newtonian dynamics, Maxwell's electromagnetic theory, the chemical revolution, Einsteinian physics, the wave theory of light, the dynamic theory of heat and quantum mechanics. (Kuhn 1962).

The currently prevailing conventional medical paradigm is often referred to as "biomedicine". Biomedicine's underlying assumptions are built upon modern scientific understandings in anatomy, biology, chemistry and physics. Ken Schaffner wrote a 2002 article exploring whether CAM healing traditions might be "paradigm-dependent". Schaffner notes how Traditional Chinese Medicine "has a quite different: (1) basic set of elements, (2) pathophysiology, (3) differential diagnosis, and (4) set of therapies" compared to biomedicine (Schaffner 2002). For example, acupuncture involves the manipulation of Qi, a vital energy, which flows along 12 specific, invisible pathways along the body called meridians. The meridian system is what most acupuncturists use to determine where on the body to place acupuncture needles.

According to Traditional Chinese Medicine theory, meridian "imbalances" are determined through symptomatology, specialized pulse and tongue diagnosis, and many other characteristics that are not thought to affect health diagnosis or treatment in conventional biomedicine. In general, CAM therapies suggest new quasi-metaphysical assumptions about what properties exist in humans and the world (biofields/vitalism/Qi), different standards of evaluation (trial and error, tradition, valuing personal subjective experiences), different kinds of data (pulse or electroacupunture diagnosis, muscle testing), different languages and meanings

("energy"), different exemplars (Samuel Hahnemann for homeopathy, Linus Pauling for vitamin C, Norman Cousins for laughter therapy) and different types of puzzles in need of solving (mind-body-spirit imbalances, intestinal yeasts, parasites and fungi). Properties of the world, standards of evaluation, types of data, languages and meanings, exemplars and types of puzzles to be solved are all the categories shared by adherents to a Kuhnian scientific paradigm, and the ways different paradigms differ (Kuhn 1962).

A 2003 article in the Journal of Law, Medicine and Ethics argues that these differences make it impossible for TCM to conduct a "stringent" RCT:

> In traditional Chinese medicine, symptom-complexes do not describe diseases. They describe the functioning of the whole body at a definite time or stage of a disease. They are differentiated according to the "eight guiding principles" (yin and yang, interior and exterior, cold and heat, deficiency and excess), the state of qi and blood, the theory of the channel, the theory of the organs (zang-fu), the etiology of disease, and so on. Accordingly, it is nearly impossible to obtain exactly the same symptom-complex for any two patients, let alone a group of patients. As symptom-complexes differ in different patients, the prescriptions of herbal medicines for treating them must differ too. Thus, it is impossible for traditional Chinese medicine to conduct a stringent randomized clinical trial. (Fan 2003)

Schaffner comes to a different conclusion than Fan on this point. Schaffner notes that an RCT of individualized acupuncture treatment is still technically feasible even if acupuncture meridian imbalances cannot be measured within a biomedical paradigm. Such a trial is feasible because the presence or absence of a medical disorder (eg chronic hives) can still be measured within a biomedical framework. Therefore, even if Qi diagnoses and optimal acupuncture points and herbal prescriptions differ in different patients, it is still possible to conduct a randomized trial where subjects with a common ailment (chronic hives) receive individualized acupuncture and herbs in one arm of the trial and placebo/sham acupuncture and herbs in another arm of a trial.

Fan may have been making the mistake where "detractors [of using RCTs for TCM] underestimate the creativity possible in accommodating Chinese medicine and RCTs to each

other. While some RCTs on Chinese medicine therapies are conducted entirely within biomedical diagnostic categories and Western constructs of the body, others strive for a design more in tune with the paradigm of Chinese medicine" (Shea 2006). Additionally, "TCM has a 'limited number of identified 'syndromes'' and patients with similar syndromes are often prescribed similar treatments" (Shea 2006). Therefore, even if different subjects with chronic hives in an RCT received individualized TCM diagnoses and treatment procedures, the different cases probably share enough similarities so that the results could be generalized to many other patients with chronic hives. Therefore the individualized nature of CAM therapies as well as some of their paradigmatic differences may not necessarily prohibit such therapies from being studied with randomized controlled clinical trials.

Schaffner does note, however, that the influence of beliefs in CAM therapy may make randomization problematic. The role of beliefs and expectations can cause the act of participating in a clinical trial to reduce the effectiveness of the therapy in a way that would not occur in clinical practice.

3.2 BELIEF PREREQUISITE

It is possible that CAM therapies like acupuncture are different than more conventional therapies in a way where they are more affected by and dependent on beliefs. I call this the "belief prerequisite" and I believe it may be stronger and more common with some CAM therapies than with other kinds of therapies, and therefore may cause more inaccuracies in attempting to subject CAM therapies to randomization in scientific testing.

The belief prerequisite creates a problem when belief in a therapy may be required for the specific effects of a therapy to be effective. Absence of belief may cause the specific effects of a therapy to have substantially less of an effect. An article by Black about observational trials and situations when randomization may be inappropriate discusses how randomization can be self-defeating and how "the very act of random allocation may reduce the effectiveness of the intervention" for "interventions for which clinicians, or patients, or both, have a preference (despite agreeing to random allocation), and where patients need to participate in the intervention--psychotherapy, for example" (Black 1996). In a placebo-controlled trial, certain subjects in the treatment group are often suspicious that they might be in the placebo treatment group, and this suspicion may substantially alter the results of the treatments.

This belief/ambiguity effect would happen in any placebo-controlled trial, but it could be argued that acupuncture and other CAM therapies are more susceptible to this kind of effect for two reasons: 1. The synergies of belief and healing may be stronger with CAM therapies because of the mind-body nature of many CAM therapies, which is not shared by most conventional therapies, and 2. because people may be less likely to believe in CAM therapies to begin with because of their unconventional natures and less-than-scientific status.

3.2.1 Synergies for beliefs and mind-body therapies

CAM therapies often address health symptoms and problems from a whole-body perspective that is difficult to separate from a mind-body perspective. By their nature their specific effects are particularly difficult to study, isolate and distinguish from non-specific effects in RCTs. I would posit that the mechanism of CAM therapies could closely resemble the mechanism of non-specific effects such as the placebo effect. "Some CAM domains could potentially reveal

theoretical models of corporeal energy fields (such as qi or prana; which are discussed in traditional Chinese and Indian medicines, respectively) that represent a 'missing link' between mind and body currently lacking in mainstream psychology" (Hughes 2008). Medical science does not currently have a robust explanation for the mechanism of the placebo effect.

If many CAM therapies are essentially targeting the same mind-body mechanism as the placebo effect, in a more specific manner, it could be argued that CAM therapies would also be more susceptible to the impact of beliefs and expectations compared to the effects of beliefs and expectations on standard medical treatments. These influences include the beliefs and expectations of the patients/subjects as well as the practitioners.

> Integrative interventions tend to involve potentially synergistic, multimodal, and complex interactions that are often dependent on the relationship between practitioner and patient, and on patients' preferences, expectations, and motivations. For example, the motivation, compliance, and response of a patient undertaking dietary or other lifestyle changes, or practicing relaxation exercises, will depend greatly on how they feel about their practitioner. Consequently, a randomized placebo controlled trial aiming to study components of integrative interventions in isolation may actually distort the very thing it is investigating (MacPherson et al 2009).

If ambiguity around real or sham acupuncture has a strong effect on an acupuncture treatment, "outcomes from trials where patient preference or treatment expectations are not considered are unrealistic estimators for the effects likely to occur in uncontrolled practice. Therefore failure to find an effect in a randomized trial cannot necessarily be taken as an indication of ineffectiveness" (Walach et al 2006). If the ambiguity in an experimental context caused patients in the experimental arm of a trial to heal less than they would in a clinical setting, and if this difference caused the trial to be negative or null/inconclusive when it would have otherwise been positive, this would be an example of a type-II error, false negative situation where an inconclusive RCT could fail to find an effect that would exist in clinical practice. The positive

effect might also exist in a non-randomized trial that specifically put people who believed in acupuncture in the acupuncture treatment arm.

The argument that randomization will distort the results of a CAM therapy trial more than in a conventional medicine trial may not be persuasive to everyone. A person might have to entertain the idea that there is something distinguishable about the role of beliefs in CAM therapies versus the role of beliefs in other therapies in order to accept this excuse for the suboptimality of randomization in CAM research. If someone did not accept that a CAM therapy might be more susceptible to beliefs and expectations because of the mind-body nature of the CAM therapy and the nature of the placebo effect, then s/he might not agree that the belief prerequisite is more applicable to CAM therapies than other kinds of treatments.

3.2.2 Disbelief may be stronger with CAM therapies

Since the purported mechanism of CAM therapies propose undiscovered properties of the human body, and sometimes even properties of the world that have not yet been discovered by science, many people have a hard time believing in their purported mechanisms. Accordingly, this might hamper the therapies' effects more often than with conventional medical treatments that are more accepted in the medical and scientific community. On the other hand, acupuncture is still known to work on some skeptical patients, so beliefs don't necessarily stop all of the effects of acupuncture. However, in a clinical trial the effects of uncertainty and disbelief in a CAM therapy across a large study population may hamper the effects of the therapy because it is averaging together the results from the therapy on a mix of believers and nonbelievers who were all randomized into the experimental arm of the trial.

3.3 PROBLEMS WITH CONTROLS AND PLACEBOS: SHAM ACUPUNCTURE ISN'T A TRUE SHAM

Another problem with CAM research has to do with placebo controls and the "efficacy paradox". For acupuncture research with placebo control arms, the placebo controls often involve "sham acupuncture" or "minimal acupuncture", a control treatment which uses non-acupuncture points and shallower needle insertion or fake needles. This is supposed to be a placebo control treatment, but it is not necessarily inert and has been debated in articles published in various medical journals.

The efficacy paradox has to do with the difference between a therapy having a statistically significant effect and a therapy being practically significant. The paradox occurs when the effects of a treatment are not found to be statistically significantly efficacious beyond a placebo control, but the overall effect of the treatment is nevertheless strongly practically effective. This effect can be seen in some large-scale three-arm German acupuncture studies conducted in the mid-2000s to test both the efficacy and effectiveness of acupuncture on lower back pain, tension headache, knee osteoarthritis and migraines.

In each study, one treatment arm was acupuncture, another arm was sham acupuncture using shallower needle insertion at non-acupuncture points, and a third arm was conventional treatment recommended by standard guidelines. "All of them tested acupuncture against sham acupuncture and the best-evidence conventional standard that was supposed to be superior and well covered by clinical trial evidence. This active comparator was not just 'treatment as usual' but the best that conventional medicine could muster at the time, according to German guidelines, designed and expected to be clearly superior to placebo and at least as effective as real acupuncture" (Walach 2011). The back pain and knee osteoarthritis studies concluded that

both acupuncture and sham acupuncture were almost twice as effective as the typical recommended treatment, which included physiotherapy and nonsteroidal anti-inflammatory drugs (NSAIDs) (Haake et al 2007, Scharf et al 2006). The migraine study concluded that real and sham acupuncture were both just as effective as the standard therapy (Diener et al 2006). However, in each trial, the large improvements in the sham and real acupuncture arms were equal, on average. Therefore, even though acupuncture and sham acupuncture were equal to or better than standard recommended conventional treatments, acupuncture did not technically have a statistically significant effect beyond sham acupuncture in the German studies, so the trials were technically null/inconclusive regarding the efficacy of acupuncture. This was technically the case even though both the real and sham acupuncture had greater *effectiveness* than the pharmacological intervention. Another similar 2009 four-arm randomized, controlled study in the United States tested acupuncture for chronic low back pain and used toothpicks in a guidetube as a placebo-like "simulated" sham acupuncture control. That study also found that there was no statistically significant difference between individualized, standardized, and "simulated" acupuncture, but that all types of acupuncture were more effective than usual care. Its conclusion was that "tailoring needling sites to each patient and penetration of the skin appear to be unimportant in eliciting therapeutic benefits" and that "it remains unclear whether acupuncture or our simulated method of acupuncture provide physiologically important stimulation or represent placebo or nonspecific effects" (Cherkin et al 2009).

One possible problem in cases like this is that sham acupuncture may not be a true sham. The distinctive nature of acupuncture appears to pose a problem for attempts to test it with a placebo-controlled trial because there may be no such thing as a true CAM sham, at least for a treatment like acupuncture. A similar argument could be made for other CAM therapies such as

biofield healing practices like Reiki and Therapeutic Touch. "In the absence of knowing how a treatment 'works,' an appropriate sham control cannot be designed rationally. And, without an appropriate sham control, the efficacy of a treatment cannot be assessed adequately" (Hammerschlag and Zwickey 2006). The incomplete scientific understanding of the mechanism of acupuncture therapy creates a barrier to conducting a proper placebo-controlled acupuncture trial. "If acupuncture points, for example, correspond to network theory's hubs and nodes of a connective tissue-based regulatory matrix, then sham needling at irrelevant sites is also likely to activate the network, albeit not quite as strongly" (Hammerschlag and Zwickey 2006). Some researchers might believe that placebo controls in placebo-controlled acupuncture trials are inert, but using acupuncture's epistemology, sham acupuncture could still be balancing a person's Qi in some way. Therefore sham acupuncture would still be having some of the same specific effects that regular acupuncture has. Ideal placebo controls are not supposed to have any specific effects. They are supposed to be inert and only have non-specific effects. According to acupuncture meridian theory, acupuncture on specific points should produce better outcomes than sham acupuncture on non-acupuncture points, but this difference may not be large or consistent enough across every individual to show up in a clinical trial. Assuming that this is true, it is not clear if it would ever be possible to have an inert placebo in a randomized placebo-controlled clinical trial of a treatment like acupuncture.

One problem with this explanation of the efficacy paradox for acupuncture is that it necessitates a basic understanding and openness to acupuncture meridian theory. An uninformed observer may think that "sham acupuncture" is a perfectly legitimate placebo acupuncture treatment as long as fake needles or non-acupuncture points on the body are used. It is difficult to design an appropriate sham control for acupuncture if the mechanism of acupuncture is still

scientifically unknown. “[A]n appropriate sham procedure cannot be designed without sufficient knowledge of what needs to be ‘shammed’” (Langevin et al 2011). It is not difficult to see how someone with no understanding of acupuncture’s Qi concept could think that putting non-needles in the wrong spots on the body might not do anything, because s/he would be assuming that the beneficial effect of acupuncture depend entirely on the mechanical details of the intervention: the types of needles and putting the needles in the proper spots. Some actual acupuncturists may not even realize that sham acupuncture can be confounding if the acupuncturists believe that it is critically important to place needles on specific acupuncture points that depend on individualized patient characteristics. These acupuncturists may underestimate the non-specific effects of acupuncture, or they may be underestimating the acupuncture-esque specific effects of sham acupuncture. Or both. Also, even if someone did understand how the concept of Qi relates to the efficacy paradox, s/he might have to be able to entertain Qi as a possible explanation for the positive effects of acupuncture. If someone were unwilling to even entertain this possibility because it was too “biologically implausible”, then it would yet again be possible for a skeptic to argue that the positive effects of both acupuncture and sham acupuncture could be due entirely to the non-specific effects of belief that are part of the placebo effect. Therefore, the skeptic could argue that acupuncture and sham acupuncture have such strong non-specific placebo effects that people’s beliefs in acupuncture are more effective for the treatment of low back pain and knee osteoarthritis than a pharmacological treatment, but that none of this has anything to do with Qi “energy” or invisible acupuncture meridians.

Even if the efficacy paradox and the sham acupuncture problem limit the accuracy of placebo-controlled acupuncture trials, it does not necessarily mean that acupuncture is incompatible with the evidence-based medicine hierarchy. The efficacy paradox only creates a

problem for a placebo control, which may be less of a problem for the evidence-based medicine hierarchy than the problems with randomization that Schaffner addresses because placebo controls are not a necessary part of the EBM hierarchy. Some critics have responded to the criticism of using placebo controls in acupuncture with the clarification that it is standard practice for "therapeutic interventions" like surgery and physical therapy (as opposed to pharmacological interventions) to get tested without placebo-controlled trials: "This is precisely the reason why the placebo-controlled, double-blind trial is not the 'gold standard' for efficacy testing of therapeutic interventions...In many areas of clinical medicine the desirable does not coincide with the achievable. The gold standard is therefore the randomized clinical trial (RCT), which is possible for all complementary/alternative treatments" (Ernst 2002). This argument is a testament to the theoretical flexibility of the evidence-based medicine hierarchy, but the conclusion here is somewhat strong: placebo-controlled trials of acupuncture could be misguided because it is impossible to have a proper placebo control, and therefore the outcomes research that is used for most surgeries might be the best way to go about studying acupuncture. "Instead of asking if acupuncture is better than placebo, a pragmatic trial inquires: "Is acupuncture of better value than what is currently on offer?" (Shea 2006). This is not a widely accepted or definitive conclusion because null/inconclusive placebo-controlled sham acupuncture RCTs are still conducted and published. Also, the acceptability of pragmatic outcomes research for acupuncture may be difficult for skeptics to accept if they think that the beneficial effects of acupuncture are due entirely to the placebo effect. Some have argued that "until acupuncture has proved its efficacy within many standard [placebo-controlled] RCTs, the results of outcomes research will mean little to many arbiters. For the paradigm of Chinese medicine to be seriously considered by skeptics, several of its therapies must first be shown efficacious within

MacPherson's 'explanatory trials'" (Shea 2006) which measure efficacy beyond placebo. This particular issue appears to be closer to a Kuhnian paradigm conflict than the paradigm conflicts addressed in Schaffner's article.

It is interesting that one of the articles that points out how non-placebo RCTs are better than placebo RCTs for "therapeutic interventions" is the same article where a common muddled presupposition is made, namely that when a CAM treatment is not shown to be better than a placebo in a trial, "these patients are obviously benefitting from a placebo effect" (Ernst 2002). This statement is muddled because it discounts false-negatives and type II errors and the possibility that certain factors could cause randomized placebo-controlled trials to be negative or null/inconclusive even if a real effect does exist or would exist in clinical practice. The article is written by Edzard Ernst, a professor in the "Department of Complementary Medicine" at the University of Exeter. Professor Ernst is a somewhat skeptical CAM researcher who has published many critical articles about CAM therapies, and in this article he even admits that the views expressed are "doubtless biased" and that he is open to criticism of his arguments. This bias may be part of the reason why he accepts the inadequacy of placebo-controlled trials for some CAM treatments in one part of the article but then appears to forget this fact later in the same article.

3.4 SUBSET ISSUE

Another possible explanation for studies that show no effect but are in fact missing a real, undetected effect (type II errors/false negatives) in CAM research is a subset/subgroup problem.

In inconclusive/negative trials there could be a subset of people who did, in fact, benefit greatly from a therapy, but this effect might not be distinguishable when averaged together and compared with the control arm of the trial, and thus the trial finds "no effect" of the therapy. "Some individuals in any large clinical trial may have causally benefited from an intervention that failed to demonstrate efficacy across the population as a whole." (Tonelli and Callahan 2001). Clinical trials with large numbers of subjects (large n) are more likely to miss distinctive effects in small subgroups because the distinctive effects get unnoticed when all the results from all of the subjects are averaged together and analyzed in tests for overall statistical significance between different arms of a trial. A study may not show an overall statistically significant difference between two arms of a trial, and so the conclusion of the study would be that the trial was null/inconclusive, even if there really was an important difference between two comparable subsets within the groups being studied. This is an inherent limitation of statistical significance testing.

This limitation is similar to the ungeneralizability problem that physicians struggle with where relying on information from seemingly conclusive RCTs may not be applicable to individual patients in the real world of clinical practice because the RCTs were conducted on a population that differs greatly from the physician's patient. In this case the problem is reversed: data from inconclusive trials may be applicable to individual patients. One hypothetical example to elucidate this problem would be a therapy that has a real effect, but the effect is normally small and smaller than any concurrent placebo effects, except for a select subgroup of subjects for which the real effect is very large. Additionally, it could be the case that, for example, some subjects who would have had an 80% improvement in a placebo control arm would instead have a 10% improvement from the therapy itself and then a concurrent 70% improvement due to

placebo effects. So the outcome (effectiveness) between the experimental and placebo control arms would be almost the same, but the treatment would still be having a small, real efficacy for every subject in the experimental arm. And efficacy is what is being tested in a placebo-controlled trial. In this case, the intervention would likely fail to demonstrate efficacy beyond placebo across the population as a whole, even if every single participant benefitted from the real, usually small, but sometimes quite large, effects of the therapy. A typical randomized placebo-controlled trial would not identify this kind of effect, which would not only be happening for every single subject but would also be having large effects for some subgroup.

One response to this criticism is that subgroup analysis would be able to solve part of this problem. If a placebo-controlled drug trial finds no statistical significance difference between a large treatment and control group, a subgroup analysis might find, for example, that comparing only the subset of study participants in each group who are over 70 years old, or who are female, finds a difference. One limitation of subgroup analysis is that the subgroup might not be large enough to find a statistically significant difference.

However, even if there are enough people in a subgroup to be able to do statistical significance calculations, subgroup analysis is only possible if a subgroup can be properly identified. "The only way to measure within-group variation is to gather the information from each patient that may mark them as a member of a relevant subgroup and then conduct analyses of these data" (Bluhm 2005). Considering the large gap in knowledge about human illnesses and the inability of medical testing to detect many medical disorders it is not hard to conceive of hidden subgroups of people who really are benefitting from treatments in "inconclusive" randomized clinical trials of large populations, or even subgroups who would benefit but none of whom happen to get included in the trial in the first place.

In the philosophy of science, the Duhem-Quine thesis argues that hypotheses are never tested in a vacuum; they always depend on auxiliary hypotheses, including assumptions about the adequacy of testing methods and basic assumptions about the way the world works (Hempel 1966). When any hypothesis is tested using the scientific method and the results of the test are negative or null/inconclusive, it is always possible that the hypothesis might still be correct and that there might have been something wrong with an auxiliary hypothesis rather than there being something wrong with the hypothesis that was being tested. Maybe the testing method was flawed, or perhaps some fundamental scientific assumption was incorrect.

The possibility of flawed auxiliary hypotheses allows any researcher to make "ad-hoc" after-the-fact defenses of his/her hypotheses to try to save the hypothesis despite a negative or inconclusive trial result.

> "[I]neffectiveness, at best, can be claimed only for the use of the particular intervention, in a particular dose or format, on a particular schedule, by a particular group of practitioners, in a particular patient population. Presumptions of ineffectiveness can be appropriately challenged by arguments that the intervention was used incorrectly, that the practitioners were unskilled, that the population studied was the wrong one. A seemingly infinite number of studies would be required to demonstrate true ineffectiveness of any particular intervention" (Tonelli and Callahan 2001).

These after-the-fact defenses are not often considered to be credible because they are after-the-fact, but a condemnation of every ad-hoc defense discounts the possibility that some ad-hoc defenses will be accurate. There is never any way to know for sure if an inconclusive test of a medical treatment means that the treatment has been invalidated or that something else was wrong with the trial. This is also known as the underdetermination of scientific theory. This concept can be used to argue that negative or null/inconclusive RCTs that appear to "disprove" a medical treatment may not be as credible and definitive as most people think they are because the averaged results of a trial might be overlooking hidden subgroups who are benefitting from

the treatment. Even a process such as the FDA clinical trial phase system, which usually requires positive randomized placebo-controlled RCTs before a drug can be sold, may be overly discounting possible beneficial effects of therapies on undetectable subgroups or possible beneficial effects that are overshadowed by the placebo effect, on average, across the subject population.

4.0 EVALUATION OF ARGUMENTS FOR WHY CAM SHOULD BE STUDIED WITH THE SAME CRITERIA AS CONVENTIONAL MEDICINE

Many researchers, physicians and academics have argued that there is nothing special about CAM therapies that would exempt them from the rigorous evidence standards used to evaluate conventional therapies. Additionally, it is thought that advances in clinical testing methods will eventually uncover creative ways to conduct high-quality research on CAM therapies which will lead the way in these disputes (Hrobjartsson and Brorson 2002). Two of the more well-known critiques along these lines were made in a special 1998 issue of the New England Journal of Medicine dedicated to the topic of complementary and alternative medicine:

> "There is no alternative medicine. There is only scientifically proven, evidence-based medicine supported by solid data or unproven medicine, for which scientific evidence is lacking. Whether a therapeutic practice is "Eastern" or "Western," is unconventional or mainstream, or involves mind-body techniques or molecular genetics is largely irrelevant except for historical purposes and cultural interest" (Fontanarosa and Lundberg 1998).

> It is time for the scientific community to stop giving alternative medicine a free ride. There cannot be two kinds of medicine — conventional and alternative. There is only medicine that has been adequately tested and medicine that has not, medicine that works and medicine that may or may not work. Once a treatment has been tested rigorously, it no longer matters whether it was considered alternative at the outset. If it is found to be reasonably safe and effective, it will be accepted. But assertions, speculation, and testimonials do not substitute for evidence. Alternative treatments should be subjected to scientific testing no less rigorous than that required for conventional treatments (Angell and Kassirer 1998).

A third critique goes as follows:

> I submit that if these [CAM] treatments cannot withstand the test of empirical research, if their effectiveness could just as easily be explained by the "natural history of disease,

> regression to the mean, suggestion, counter-irritation, distraction, expectation, consensus, the Stockholm effect (identifying with and aiding the desires of a dominant figure), fatigue, habituation, ritual, reinforcement, and other well-known psychological mechanisms," then we have wasted a lot of time and effort. The time has been wasted on all the people who have spent years learning falsehoods about acupuncture points and the principles of homeopathy. And the patients have wasted their time, money, and efforts receiving treatments that were not what they were represented to be or were harmful… how can we make rationing judgments in the absence of good data? As with all of medicine, we should test these practices by the same measure as we test medicine itself (Schneiderman 2003).

I have four primary criticisms of these arguments for how rigorous evidence criteria should be applied to CAM. One is that these arguments set a double standard that conventional medicine itself does not usually meet. Most of the practice of medicine does not meet the gold standard criteria of RCTs and meta-analyses. A second criticism is that the gold standard in the EBM hierarchy is not as useful as it is perceived to be. A third criticism is that interpreting evidence for CAM therapies may be paradigm-dependent. And a fourth criticism is that economic, cultural, and sociological factors play a large role in what therapies are accepted into conventional medicine. Some therapies are not studied because there are not enough financial incentives to do so, but even when there is evidence and proof to support unconventional therapies they are still not readily accepted because the evidence and proof is not considered good enough or because people choose not to incorporate or recommend them regardless of how much supportive evidence there is for them.

4.1 DOUBLE STANDARD

4.1.1 Much of conventional medicine is by and large not evidence-based

The three quotations above hold CAM to a double standard because much of modern conventional medical practice is not evidence-based or rigorously tested. The real-life practice of medicine in actual clinical settings "does not and never can measure up to the scientific standard to which critics of CAM would like to hold alternative modalities" (Morreim 2003). A few estimates have been made since the 1980s that only 10-20% of clinical interventions are supported by studies and are therefore "evidence-based". A later study from 2000 estimated that an average of 76% of interventions have "compelling" evidence to support them while 37% of interventions are "evidence-based" in terms of being supported by randomized evidence (Imrie and Ramey 2000, Howick 2011).

4.1.2 Off-label prescribing

Off-label prescribing involves physicians prescribing drugs to patients to treat certain medical conditions even though the drug has only been specifically approved by the FDA to treat other medical conditions. For example, Neurontin, a drug approved to treat epilepsy, can be prescribed by a physician to treat pain or anxiety. The commonly accepted practice of off-label prescribing of pharmaceuticals is unsystematic and often unsupported by rigorous evidence, or any clinical studies at all.

> By some estimates, as many as one-half of all prescriptions are for offlabel uses. An AMA study estimated that 40% to 60% of prescription drugs were given for unapproved

> uses. According to one survey, offlabel prescribing of top-selling drugs *doubled* from 1998 to 2003. Another study indicated that most hospital patients receive at least one drug off label (Johns 2006).

Off-label prescribing is not universally accepted in every case. The FDA has warned that "under-evaluated off-label practices-those not supported by strong scientific evidence-may 'jeopardize patient safety or represent economically wasteful prescribing practices.'" (Johns 2006). Jeopardized safety and waste are potential problems that can be caused by off-label prescribing, but it should be noted that medical care that is unsystematic and unsupported by evidence is not necessarily bad medicine. "[F]or many years aspirin was prescribed to reduce the risk of heart attack but was not approved for this use until 1998…Moreover, off-label uses are frequently the *only* available treatment for some serious and fatal diseases. Specifically, 62% of cancer patients use drugs off label. Indeed, off-label uses in some cases represent the highest standard of care, especially in treating AIDS, cancer, and rare diseases" (Johns 2006). So off-label prescribing is not always evidence-based, but it is also not always a bad thing to do. "Of course we want everything to be evidence-based, said Dr. Yashar Hirshaut, an oncologist in Manhattan. I also like the American flag and apple pie. But, he explained, you say, 'This person is dying right here and I need something that will help, and there's a logical construct that I can see how it will help'" (Kolata and Pollack 2008). In an ideal world, all pharmaceuticals would be tested for safety and efficacy for all the types of conditions for which doctors prescribe them. However, in the absence of such expensive and impractical informational certainty, most physicians are comfortable prescribing drugs to treat their patients' medical conditions in a way that would fit in at the bottom of the EBM hierarchy.

4.1.3 Many unambiguously effective medical interventions have never been subjected to randomized, controlled trials

Many medical treatments regarded as unambiguously effective have never been subjected to randomized controlled trials. Observational trials from the middle of the hierarchy are often the best forms of evidence for many widely accepted treatments. “It is most unlikely that any ethics committee in an industrialized country would sanction the random allocation of patients to intensive care versus ward care, or cardiac transplantation versus medical management.” (Black 1996). Diuretics in heart failure, external defibrillation to start a stopped heart, tracheostomy to open a blocked air passage, the Heimlich maneuver to dislodge airway obstructions, and many surgeries “were never (and now, let us hope, never will be) ‘validated’ in an RCT” (Worrall 2007). The reason these procedures will likely never be validated in an RCT is because conducting such a trial would likely result in unnecessary deaths in the control arm of the study. Furthermore, the benefits of the procedures are already considered proven enough through common practice and observation. That is why they are considered unambiguously effective.

4.1.4 Most surgeries have never been studied with placebo-controlled trials

Modern surgery functions with much success without subjecting many of its practices to one of the most rigorous forms of evidence, randomized placebo-controlled trials. Most surgeries have never been subjected to placebo-controlled trials because a placebo treatment for surgery involves cutting people open and sewing them back up without performing the actual surgical intervention inside the body. For most surgical procedures it is considered unethical to conduct

placebo controlled trials, and such testing is not mandated by regulatory authorities (Wartolowska et al 2014).

Some placebo-controlled surgery trials have been conducted because there was enough ambiguity over whether the benefits of those particular surgeries were greater than placebo effects. This has been done for arthroscopic surgery for knee osteoarthritis and vertebroplasty for lower back pain, among others. The results of those sham surgery trials were null/inconclusive, indicating that some or all of the benefits of such procedures could be attributed to non-specific effects and placebo effects rather than the specific effects of the surgeries themselves (Wartolowska et al 2014).

The important point here is that most surgeries do not have the most rigorous kinds of evidence to support them, and are still accepted. Therefore it cannot be claimed that CAM therapies needs to be subjected to the highest scientific standards in order to be accepted if a large subset of medicine like surgery does not meet the same highest standards. The idea that most conventional medicine is evidence-based and has the most rigorous standards for its practices is more of an ideal than a reality.

4.1.5 Other complex interventions are not studied with placebo controls

Surgery is considered a complex intervention, along with psychotherapy, physical therapy/physiotherapy, palliative care, etc. Because of methodological and ethical reasons, the majority of these kinds of therapies are considered to be exempt from having randomized placebo-controlled studies conducted on them. This has not stopped these treatments from being performed, though there is some scientific ambiguity associated with physical therapy and psychotherapy for the same reason there is ambiguity of acupuncture research: it is difficult to

separate the specific effects from the non-specific effects (Paterson and Dieppe 2005). There is less perceived ambiguity over surgery. And there seems to be more perceived ambiguity over CAM therapies that could also be considered complex interventions. The perceived ambiguity over complex interventions that are not normally subjected to placebo controls might even be placed on a continuum, with CAM having the most ambiguity, then psychotherapy, then physiotherapy, and then surgery.

4.1.6 Antidepressant research does not meet the highest scientific standards

Antidepressant research is particularly unrigorous. Many placebo controlled antidepressant trials are null/inconclusive or find very small differences between treatment and placebo control groups. Some have argued that there is no real evidence that antidepressants "do anything", and that their effects are due exclusively to beliefs stemming from the fact that people experience side effects and assume that they are working (Angell 2011).

This is similar to the arguments made against CAM therapies. With CAM therapies, the ritualistic aspects may supply a similar magnitude of placebo effects as the placebo effects that occur from the presence of side effects in antidepressant trials. However, because of the subset issue and other issues with CAM research discussed above, along with other methodological issues related to observational studies and anecdotes, I do not believe that it is a sound argument to conclude that, since rigorous placebo-controlled trials tend to be null/inconclusive, antidepressants, or CAM therapies, don't "do anything". There are other evidential factors to consider in the debate about whether a therapy "works" in addition to the results of placebo-controlled trials.

4.1.7 Other reasons for double standard

There are other parts of conventional medicine where accepted practices are not scientific or entirely supported by rigorous evidence. In 2003 a GlaxoSmithKline executive publicly highlighted a 2001 study to note that 90% of drugs work for only 30-50% of people who are taking them (Connor 2003, Spear et al 2001). This means that many patients are exposed to side effects of drugs that aren't even working for them in the first place (Pollack 2008). Unfortunately, not enough is known about why some patients respond to certain drugs and other patients don't, or which patients will experience side effects, or why some patients experience some side effects and other patients experience other side effects.

Long-term safety of many pharmaceutical drugs is often determined through long-term observation of patient use rather than through long-term clinical trials. Long-term patient observation is a less-than-rigorous way of gathering evidence. It involves noticing case reports and patterns of adverse reactions to drugs over longer time periods than the time periods the drugs had been studied. The use of case reports and noticing patterns is the same way that many CAM providers evaluate their treatments (Snyder 2007).

Medical insurance does not require therapies to be supported by the most rigorous types of evidence. The standard that insurance companies use is often referred to as medical necessity.

> Insurers generally evaluate whether a treatment is medically necessary by looking to the customary practice within the medical community. Customary practice among medical doctors or physicians, however, is often the result of trial and error. Doctors' prescription and treatment patterns are based frequently on their own clinical experiences - not on studies proving that the particular treatment is effective. Studies for many treatments simply do not exist. Moreover, existing studies may reach conflicting results, like those regarding the effectiveness of mammograms and hormone replacement therapy (Atwell 2003).

Standard of care is also a consideration that physicians keep in mind when they practice "defensive medicine" to defend themselves from malpractice lawsuits (Bishop et al 2010). The standard of care in medicine is a legal term that refers to the level of medical care that can reasonably be expected from a skilled practitioner. It is the standard that health care providers would need to meet in order to avoid negligence. Standard of care is also not based on the most rigorous studies; it can be based on customary practice in the entire medical community or even in a specific region or other subset of the medical community (Timmermans and Berg 2003). Many physicians practice medicine in a way that is focused on how they could be sued for malpractice if something is missed or goes wrong. This leads to overtreatment and unnecessary testing and imaging by physicians who are trying to cover their bases and avoid potential liability. In theory this is not good medicine and should not happen as often as it does, but it is common and a large factor influencing contemporary healthcare practice.

4.1.8 Ideal vs reality

An important issue in these debates is the way people think of RCTs and EBM as if they function in their ideal forms, when they never in fact reach that ideal. EBM supporters have too idealized a notion of that which they support. "At least some advocates for evidence-based medicine seem to be praising the RCT methodology based on what it is hypothetically capable of and then playing a sort of bait-and-switch by providing us with the results of real RCTs as if they held the answer to all clinical questions" (Brody 2002). Medical science is more of a *pursuit* of rigorousness and unattainable certainty, but it is often perceived as having achieved much more rigor and certainly than it actually ever achieves. There is a "vanity of rigor in RCTs" (Brandt 1990).

4.1.9 Replication Issues in Medicine

In the ideal of scientific medicine, studies get replicated and validated over time. However, people tend to underestimate how uncommon it is for studies to get replicated, and how often it is that replication studies find different results from an original pathbreaking study. Many important medical studies that find large effects do not find the same magnitude of effects when they are replicated in the future (Ioannidis 2005). According to *Nature*, a third of all scientific studies never even get cited, let alone replicated. (Lehrer 2010).

4.1.10 The practice of evidence-based medicine is not evidence-based

In some specific circumstances, evidence-based practice appears to be more beneficial than relying on clinical intuition. However, in many other situations there is no evidence that practicing medicine according to the accepted evidence-based medicine criteria would provide better health care on average. "A fundamental assumption of the EBM movement is that practitioners whose practice is based on an understanding of evidence from applied health care research will provide superior patient care compared with practitioners who rely on understanding of basic mechanisms and their own clinical experience. So far, no convincing direct evidence exists that shows that this assumption is correct" (Haynes 2002). There are surely many cases where more and better evidence leads to safer and more effective health outcomes. For example, in 2013 the American Academy of Pediatrics released an evidence-based clinical guideline on the diagnosis and management of childhood ear infections. Their new guidelines, based on more recent high quality evidence, revised an older 2004 guideline and added a

recommended option of observation with close follow-up and deferment of antibiotic treatment for 2-3 days (Lieberthal et al 2013). Physicians who adopt this practice will likely avoid situations where they would have otherwise prescribed antibiotics unnecessarily. Therefore following these evidence-based guidelines will reduce unnecessary antibiotic side effects in patients in addition to reducing the chances of some patients developing antibiotic resistance in the future. However, even if there are some specific cases where evidence-based practice has benefits, it is not clear whether practicing medicine according to the criteria of the evidence-based medicine movement would live up to the promise given to it by its proponents if it were more widely practiced across every single aspect of healthcare.

In fact, evidence-based medicine movement could even be considered a step backwards in attempts to treat or solve some conditions if the EBM framework focuses clinical medicine too heavily towards solutions that are the most rigorously proven. A greater focus on following the criteria of evidence-based medicine could also require physicians to spend more time reading new research articles, which could take time away from potentially more beneficial duties. Considering all of the remaining uncertainty in medicine, a greater focus on evidence might not even make much of a difference in clinical outcomes for some patients. "Many conventional physicians ignore or fail fully to assess the evidence available from RCTs and nonetheless practice very effectively using broadly anecdotal or other empirical criteria, a charge often leveled at practitioners of complementary medicine" (Whitmarsh 2002). Once of the driving forces of the EBM movement is that physicians may not be practicing effectively enough with anecdotal and other empirical data, but there is no solid proof of how much outcomes would improve if medicine were that much more evidence-based overall.

4.1.11 Additional Biases in Clinical Trials

Reducing bias is not all about attempts to increase methodological rigor with randomization, control groups, placebos, inclusion and exclusion criteria, and blinding. Bias can occur in medical research from many other avenues.

An oft-cited issue is that scientific studies suffer from publication bias and the "file drawer problem", where negative or null/inconclusive studies are left in the figurative file drawer and never published. As many as half of all clinical trials are never published (Wolford, 2014, Riveros et al, 2013). Scientists and journals tend to prefer to only publish positive results over null results, especially if the scientists' funding, reputation, or relationship with special interests depends on positive results.

There is a documented history of pharmaceutical companies neglecting to publish studies that would have cast doubt on the effectiveness or safety of certain drugs. Examples that have come to light include Janssen Pharmaceutica not publishing clinical trial results indicating deadly side effects from the antiarrhythmic drug lorcainide in the 1980s, GlaxoSmithKilne not publishing clinical trial results in the late 80s and early 90s indicating questionable efficacy and increased suicidal tendencies in children and teens who took Paxil, an SSRI antidepressant, and Roche not publishing clinical trials and later refusing to share trial data indicating questionable efficacy and psychiatric side effects of the antiviral influenza drug Tamiflu in the 2000s (Gøtzsche, 2012).

There have been some strides to address this issue, including the major scientific journals requiring that all studies get preregistered in order for them to get published. However, any meta-analysis of trials conducted before the 2000s will be affected by this prior publication bias

(Jensen 2007). Furthermore, these strides cannot address all scientific bias issues. Other biases and manipulations that occur in scientific studies are the following:

- Choice of variables that are and are not measured. (Wendland 2007)
- Focus on statistical significance rather than practical significance (Ziliak and McCloskey 2008). "If you can show there is a statistically significant difference between a drug and placebo or a comparator, that's taken to mean the drug is better—even if the size of that difference is paltry." (Kendall and Cochrane 2007)
- Positive outcomes are more frequently reported for privately funded drug trials even when the quality of studies appear to be the same and methodological rigor is consistent between funded and non-funded trials (Borgerson 2009)
- Design or data manipulations:
    - Suboptimal dosing of the competitor's drug in a head-to-head trial (Borgerson 2009)
    - Publication of only part of the results of a trial (Borgerson 2009)
    - Analysis on the basis of secondary endpoints when primary endpoints do not indicate a significant effect of the treatment (Borgerson 2009)
    - Selective inclusion criteria that "engineers out" the possibility of adverse events being seen (Petryna 2007)
    - Make one drug look superior by "engineering up" a side effect in another drug by doubling the dose of the other drug (Petryna 2007)
    - Selectively including or excluding patients who had withdrawn due to side-effects in the analysis (Howick 2011)
    - Purposeful and/or erroneous calculation errors to make a treatment appear more effective than a control when in fact it is not, which happened with some NSAID trials (Howick 2011)

4.1.12 "Same rigorous standards of good science"

The White House Commission, a strong CAM advocate, expressed concern that recognition of and demand for treatment should not preempt a rigorous standard: "[M]ost CAM modalities have not yet been scientifically studied and found to be safe and effective. The fact that many Americans are using CAM modalities should not be confused with the fact that most of these modalities remain unproven by high-quality clinical studies. The Commission believes that conventional and CAM systems of health and healing should be held to the same rigorous

standards of good science" (FINAL REPORT of the White House Commission on Complementary and Alternative Medicine Policy 2002). Given the topics covered in this section, it is not clear if the conventional system of health and healing is in fact held to a rigorous standard of good science. Or, if it is, then many CAM practices may have already met that standard.

4.2 METHODOLOGICAL LIMITATIONS OF THE GOLD STANDARD IN THE EVIDENCE-BASED MEDICINE HIERARCHY

The quotations at the beginning of this section suggest that CAM ought to meet the same high standards of scientific medicine. They mention "good data", "solid data" and rigorous testing. Many consider rigorous testing to be RCTs and meta-analyses, the gold standard on the evidence-based medicine hierarchy. However, there are some inherent limitations to the common perception of gold standard evidence and randomized, controlled trials being universally superior to lower forms of evidence.

4.2.1 Philosophy of Medicine critiques of randomization

Howick and Worrall, among others, have argued how randomization is not as perfect as people think it is. Randomization can be highly valued in a way that confuses the pursuit of rigor with the actual attainment of it. An overview of the conventional view of randomization is as follows:

Kunz and Oxman (1998) declare that: "Randomisation is the only means of controlling for unknown and unmeasured differences between comparison groups as well as those that are known and measured" And therefore, as stated succinctly by Pocock and Elbourne (2000): "Only

randomized treatment assignment can provide a reliably unbiased estimate of treatment effects" (Grossman and Mackenzie 2005).

Two groups in a study would have to be perfectly balanced in all respects for an estimate of treatment effects to be entirely unbiased, and this never truly happens in reality (Bluhm 2009). Randomization is not as reliable as it appears to be. "The assumption that randomization always reduces bias creates what Horwitz (1987) terms 'the illusion of homogeneity,' which results in a failure to consider the numerous other factors that can go wrong" (Grossman and Mackenzie 2005). Borgerson argues that "research methods ranked highest in the hierarchy provide no greater guarantee that biases have been minimized than those below" (Borgerson 2009) and Worrall argues that, since randomization cannot ever randomize for all factors, and since randomization can only make two groups "*probably* equal in all other respects" (Worrall 2007), it should not be given the scientific prestige that it tends to receive.

Howick has argued that randomization would tend to reduce bias better than anything else, but he also notes that it is not clear that randomized trials will be superior to observational trials in all instances (Howick 2011). There could be many instances where lower hierarchy types of evidence are superior to higher hierarchy types of evidence, at least some of the time, and especially if randomization is nullified by one of many possible factors.

> Just as being tall is often a good property for a jumper to have, the property of being an RCT is often a good property for a study to have, but it does not follow that anything that is an RCT is better than anything that isn't. When two dissimilar studies are being compared, as is the case in practically all funding decisions, whether the studies are RCTs is not the only question to ask. It cannot be assumed that the RCT is reliably at least as good as the alternatives in *every* situation in which it is available, and it therefore follows that the RCT is not a gold standard (Grossman and Mackenzie 2005).

One example of when an RCT is inferior to an observational trial is a study on the effectiveness of botox. "Bakheit (2004) outlines the specific case of the effectiveness of botulinum toxin treatment on post-stroke muscle spasticity, which demonstrates that RCTs are less suitable, due

to their inferior ability to address 'the impact of the intervention on functional abilities and social participation'" (Grossman and Mackenzie 2005). This may also be an example where a trial is affected by the choice of which variables are and are not measured (Wendland 2007).

4.2.2 What about systematic reviews and meta-analyses vs just RCTs?

Systematic reviews are systematized ways of choosing relevant medical literature to address a certain topic, and meta-analyses are statistical approaches of combining data derived from systematic reviews.

Meta-analyses are supposed to be a way to protect from bias inherent in many RCTs, especially the kind of bias caused by small population sizes (in statistics language, small n) from single trials. However, given all the inherent biases in RCTs that have been discussed, meta-analyses are not "doing anything more than pooling the biases of individual studies, and—crucially— failing to acknowledge these biases in the end product, whether meta-analysis or guideline" (Borgerson 2009). So meta-analyses are pooling not-completely-unbiased trials together and then not fully or reliably acknowledging the possible biases. "The users of evidence are further removed from the data (of all types), and thus they are less able to critically evaluate that data for biases" (Borgerson 2009).

In addition to the problems with biases in the actual studies that are being pooled together, another issue that makes meta-analyses less reliable are the choices made about which studies to include. Inclusion and exclusion criteria for meta-analyses vary widely. "[T]his…is something of a dark art: complex protocols (which often differ at least in part from account to account) for how to rank and combine and how to produce some overall 'result' concerning the efficacy of the treatment concerned are laid down, but their underlying rationale is unclear"

(Worrall 2007). Many meta-analyses will only include RCTs, or only blinded placebo-controlled RCTs. As discussed, this attempt to reduce bias, this pursuit of rigor, is not as uniformly reliable as most people perceive it to be. The practice of relying exclusively on RCTs for systematic review excludes lower-on-the-hierarchy trials that may have equal or better evidential value than the RCTs being included in the meta-analysis.

4.2.3 N of 1 Trials

A few EBM hierarchies and EBM philosophers have discussed the N of 1 trial as a non-RCT type of trial that deserves a high spot on the EBM hierarchy and significant epistemic weight (Guyatt et al 2008). N of 1 trials are trials where a therapy is given to a single person, then taken away, and then given again. "Demonstrating causality in a single individual is most convincing when long-standing signs or symptoms abate shortly after the institution of a therapy and recur in the absence of that therapy, only to disappear once again when therapy is re-instituted" (Tonelli and Callahan 2001). The main drawback of N of 1 trials is that they are only applicable to the subject of the trial, and are therefore not generalizable to anyone else. However, they have been suggested as ways to avoid the problems that occur when conducting large RCTs. Therefore they can be "suitable for settings which do not lend themselves easily to large scale studies (e.g. the practice of some complementary practitioners)" (Ernst 1998).

N of 1 trials can be useful, but they may not be the perfect solution for some CAM therapies. CAM therapies often claim that they target the underlying cause of some medical disorders, especially chronic disorders. N of 1 trials arguably work better on therapies that target symptoms. If you take away an effective therapy that treats symptoms, the symptoms will come back, and then you can reintroduce the therapy to take them away again. If CAM therapies treat

the underlying cause of some disorder, then the disorder may not revert back to a previous heightened state after the therapy is temporarily removed. The disorder may remain slightly improved, and therefore make it difficult to separate the CAM therapy from any non-specific effects and placebo effects.

4.2.4 Large Effects make randomization unnecessary

As discussed, many EBM advocates and commentators on clinical evidence perceive randomization as a critically necessary trait for a trial to be rigorous and unbiased. However, some theorists have discussed how clinical observations can be so dramatic that randomized trials may be unnecessary. Even David Sackett, one of the pioneers of the Evidence-Based Medicine movement, recommended looking for "whether the treatment effect is so huge that you can't imagine it could be a false-positive study" (Sackett 2000) when there are no randomized studies available. In many cases it is not necessary to have a randomized trial to have solid evidence to support the effectiveness of a treatment.

> When the effect of an intervention is dramatic, the likelihood of unknown confounding factors being important is so small that they can be ignored. There are many well known examples of such interventions: penicillin for bacterial infections; smallpox vaccination; thyroxine in hypothyroidism; vitamin B12 replacement; insulin in insulin dependent diabetes; anaesthesia for surgical operations; immobilisation of fractured bones. In all these examples observational studies were adequate to demonstrate effectiveness (Black 1996).

Dramatic effects may be a promising avenue for CAM research. Due to the difficulties in conducting rigorous research on CAM therapies, dramatic observational trials may be a better method to support certain therapies and convince skeptics. I think of this as the Jackie Robinson strategy; no one could deny Jackie Robinson's talent and professionalism so he was able to break the color barrier in baseball despite societal resistance. One obstacle to this strategy for CAM

does, however, hinge on mechanism and biological plausibility; no matter how dramatic an effect a therapy has, doubters could still try to claim that the dramatic effects are due to dramatic placebo effects or other non-specific effects.

4.3 PARADIGM ISSUES

Another problem with the argument that alternative medicine should be held to the same high evidence standards as conventional medicine is that, as mentioned in the Schaffner paper, interpretation of evidence about CAM therapies may be paradigm-dependant.

> Many of the therapeutic effects of CAM modalities that laypeople seek and value are not recognized as valid in the conventional paradigm; others are defined as 'nonspecific,' rendering them unattributable to particular modalities or interventions, or nonamenable to study using typical quantitative research designs and extant scientific standards. Examples include subjectively noted improvements in general vigor or sense of well-being, promotion of active wellness, support of immune system or of an inherent recuperative capacity of the body, restoration of internal states of balance or harmony, spiritual benefits, or restoration of proper flow and function of qi or other vital energies. This disjunction poses a special challenge to CAM research, one that should not be overlooked (O'Connor 2002).

Non-placebo, non-randomized, non-blinded outcomes studies of CAM therapies may never convince skeptics in the biomedical paradigm who think that a therapy is biologically implausible, and arguments for why the nature of CAM therapies cannot be blinded, randomized or placebo-controlled may not be convincing enough to skeptics.

Interpretation of inconclusive sham acupuncture trials may depend on one's openness to the plausibility of Traditional Chinese Medicine acupuncture theory (meridians, Qi), which resembles a conflict in scientific paradigms. In fact, many CAM therapies may fit what Thomas Kuhn describes as pre-paradigm theories, where some anomalies in the predominant paradigm

are accounted for in the new paradigm, but the theory behind the new paradigm is still in a disorganized state and in need of further clarification. In this case, anomalies would be any apparent therapeutic successes after the use of a complementary and alternative medicine therapy which cannot be easily attributed to non-specific effects. Any documented phenomena that corroborate the theories behind some CAM practices could also be considered anomalies that cannot be explained within the biomedical paradigm, and may be better accounted for in a CAM paradigm.

There is some ambiguity about what Fontanarosa and Lundberg might consider “solid data”, “proven medicine” or “scientific evidence,” or what Angell and Kassirer might consider “adequate” or “rigorous” testing for complementary and alternative medicine practices. Whether a scientist considers data to be solid or testing to be adequate or rigorous could well depend on the scientist’s fundamental assumptions within his/her scientific paradigm. Marcia Angell has even admitted this point. At a 1999 conference she claimed that,

> to be good science a study must offer a "plausible biological mechanism" for the effects reported. Otherwise, the study would not be believable. She said that therapeutic touch, homeopathy, moxibustion, and intercessory prayer are examples of practices that are "preposterous" and "impossible" because they lack a plausible biological mechanism. Additionally she claimed that studies of these practices are only being published for social and political reasons. These remarks extended those that she has made in print. Similar comments were made by some of the other editors on the panel. (Hufford 2003)

Dr. Angell’s defense of biological plausibility implies that she might never consider a therapy like homeopathy to be “effective” regardless of an empirical result from any kind of trial. Many physicians have contested positive results from placebo-controlled homeopathy trials. “On the one hand, randomized controlled trials of complementary therapies that are negative tend to convince the skeptical conventional physician of ineffectiveness. On the other hand, it seems that positive randomized controlled trials or analyses in complementary therapies do not seem to

convince of anything and generate much adverse comment in the conventional medical world" (Whitmarsh 2002). One commentator, Simon Wessley, an authority on mind-body conditions, suggests that every positive trial of homeopathy is an example of an inaccurate RCT and a failure of randomization simply because it is finding a positive result of a treatment that he is automatically assuming has no benefit (Wessley 2007). If placebo-controlled homeopathy trials are not accepted, it might be an even tougher sell to argue that randomized controlled trials are not as optimal as pragmatic comparative effectiveness outcomes research for homeopathy within the evidence-based medicine hierarchy. Therefore judgments about what would be considered the "current best evidence in making decisions about the care of individual patients" (Sackett 1996) according to the principles of the evidence-based medicine movement, could be paradigm-dependent.

In addition to comments about "solid data", Fontanarosa and Lundberg contend that the specific nature of a therapy is "largely irrelevant" for scientific testing (Fontanarosa and Lundberg 1998). However, the nature of a therapy may actually be quite relevant to determining which kind of methods from the evidence-based medicine hierarchy should be used to study it because of the belief prerequisite and problems with sham controls. Accepting non-placebo-controlled research on acupuncture might require an openness to the epistemology of Traditional Chinese Medicine, or at least an ability to tolerate the uncertainty surrounding the mechanism of acupuncture and other CAM therapies. Also, anyone designing a trial for acupuncture might have to understand acupuncture theory in order to be sure that any hypothesis that is being tested is actually something that practicing acupuncturists claim and believe. Yet the criteria of evidence-based medicine might even favor CAM in this regard because "RCTs are an admirably pragmatist methodology, in the metaphysical and epistemological senses of the term" (Ashcroft

2004). This is the point that Fontanarosa and Lundberg were trying to make: the specific characteristics of any therapy should be irrelevant if the therapy is supported by solid data. However, if a health professional presupposes that any positive health effects from a therapy like homeopathy would always be due to the placebo effect, then s/he would probably not consider non-placebo-controlled research to be solid data. In this kind of situation, if an unconventional therapy achieved an outcome superior to the best available conventional treatment for a particular condition, the benefit might simply be written off as a variant of voodoo, faith healing, the power of suggestion or spontaneous remission. On the other hand, superior outcomes research could eventually persuade a skeptic to reevaluate his/her skepticism.

This tension can be seen in a 1996 Los Angeles Times article about qigoing, the Chinese "energy" balancing exercise therapy: "Doctors scoffed at qigoing, said we could put Terry in a corner and throw chicken blood on him if he thinks it's helping. Now they're all impressed with Terry's progress. But I've stopped caring what they think" (Glionna 1996). None of these issues imply that CAM therapies cannot be studied with some version of the criteria of evidence-based medicine. A number of CAM proponents support the application of the criteria to CAM therapies. The CAM conundrum may have less to do with the therapies themselves. Instead, the conundrum may have more to do with the last part of the LA Times quotation which indicates how proponents of unconventional therapies might not care about what scientists think. Even if unconventional therapies can be studied with the criteria of evidence-based medicine, the problematic attitude of the EBM hierarchy may create too many obstacles for unconventional medicine to get studied in the first place.

4.3.1 Herbalism Example

It might seem that herbal medicine could fit the biomedical pharmaceutical model of placebo-controlled RCTs more than other CAM therapies. Yet even with herbalism there are methodological issues and paradigmatic considerations. “Although often biochemically plausible, such [herbal] therapies are occasionally proffered on the basis of theories that bypass formal biochemistry. For example, some herbalists argue that their remedies do not act like ordinary drugs, but rather serve to harmonize the body's vital energy and life forces. Some therapists individualize their treatments so that patients with similar symptoms might be prescribed different remedies. Many herbal therapists prescribe multiple herbs in complex combinations” (Hughes 2008). Little evidence exists to support these styles of complex herbal prescription practices.

Even if herbs are proffered on the basis of formal biochemistry rather than theories about vital energy, it may still be problematic to study them in standard RCTs the same way that drugs are studied because of the possible synergistic aspects of many herbs. There is some evidence to support the synergistic aspects of mixtures of herbs. Sometimes this synergy occurs when the phytochemicals in the herbs are all active, but it can even occur when they are inactive and their inclusion in a mixture is still seen to increase the observed effects. “[C]omponents of plants that are not active themselves can act to improve the stability, solubility, bioavailability or half-life of the active components. Hence a particular chemical might in pure form have only a fraction of the pharmacological activity that it has in its plant matrix” (Bone and Mills 2013). Because of these synergies, neither a biochemical investigation nor a placebo-controlled efficacy trial of a single herb might capture its possible beneficial effects. A trial of herbal combinations might be possible, but if the combination is individualized to each patient, then it might only be

informative to conduct a pragmatic trial that allows for individualized combinations under the herbalists' paradigm rather than an efficacy trial that is more rigorous and tests the same combination on people with the same symptoms.

4.4 SOCIOLOGICAL, CULTURAL, AND ECONOMICS ISSUES: PROBLEMS WITH THE ATTITUDE AND PRIORITIES OF THE EVIDENCE-BASED MEDICINE FRAMEWORK

The evidence-based medicine hierarchy may be technically flexible enough to accommodate CAM therapies, but the attitude and priorities of the EBM hierarchy are still problematic if placebo-controlled RCT trials for CAM are less informative than pragmatic outcomes and comparative effectiveness trials due to Type-II false negative error problems with placebos controls, randomization and blinding.

In theory the framework of evidence-based medicine can accommodate CAM therapies, but flexibility does not matter if studies are not performed or are not performed adequately due to disagreements over whether CAM therapies deserve EBM's elasticity. The evidence-based medicine hierarchy values RCTs above all else. "Clinical and health services research communities have come to accept hierarchies of evidence where RCTs are considered the highest level of evidence and anything less than RCT-level evidence is considered somewhat suspect." (Horn et al 2005). So far this paper has outlined various reasons for why it may be problematic to attempt to study CAM therapies with placebo-controlled trials and even non-placebo RCTs. If the belief prerequisite and sham CAM problems are taken into consideration, pragmatic comparative effectiveness research may be a more optimal way to study

unconventional therapies. However, the optimality of pragmatic trials runs somewhat counter to the EBM value system. Even if studying CAM with pragmatic trials does not technically violate the evidence-based medicine framework, it violates some of the attitude and priorities of the evidence-based medicine movement. "If we make one type of evidence the 'gold standard' and orient our research approaches toward that approach, we…assume that type of information is the only valid goal for research to pursue" (Jonas 2002). It is unlikely that pragmatic research will settle all of CAM's research issues. If certain CAM modalities are classified as complex "therapeutic interventions" in the same way surgery and psychotherapy as classified as such, then pragmatic research may become one adequate way to study CAM therapies. Pragmatic research may also be compatible with the NIH mandate to conduct more comparative effectiveness research and the goals of the Patient-Centered Outcomes Research Initiative (PCORI). Yet outcomes testing will never be able to verify the underlying theories of CAM therapies or their proposed mechanisms.

Then again, positive RCTs and placebo controls for CAM therapies might not be able to verify theories and mechanisms either. Therefore, in addition to pragmatic RCTs, even less rigorous observational methods might play a role in gathering potentially valuable information that unconventional therapies, CAM or otherwise, could contribute towards basic questions about health and disease. Observational case studies also lie within the flexibility of the evidence-based medicine hierarchy and they are a subset of comparative effectiveness research, but they are even more at odds with the EBM attitude and even lower on the EBM hierarchy.

4.4.1 Observational Studies

Partly because of their low position on the hierarchy, observational studies are not always considered informative for conventional therapies, let alone for CAM therapies. "In his summing up at a major conference held in 1993, the eminent medical epidemiologist Richard Doll concluded that observational methods "provide no useful means of assessing the value of a therapy" (Black 1996). Conventional practitioners could risk their reputations by publishing CAM observational trials or CAM case studies, and many CAM practitioners are unfamiliar with the entire scientific testing and publishing process. There are different kinds of observational evidence that can serve different purposes and have different evidential weights. Some researchers such as Jeremy Howick and Tricia Greenhalgh have tried to create updated hierarchies that take into account these factors (Howick et al 2011), but it is not clear if their efforts have been very influential. "The widely held view that experimental methods (randomised controlled trials) are the 'gold standard' for evaluation has led to the denigration of non-experimental methods, to the extent that research funding bodies and journal editors automatically reject them" (Black 1996). It doesn't seem to matter that "The results of well-designed observational studies . . . do not systematically overestimate the magnitude of the effects of treatment as compared with those in randomized, controlled trials on the same topic" (Concato et al 2000). One possible avenue for the further acceptance of CAM therapies may be the use of narratives for evidence discovery and translation (Meisel 2011).

4.4.2 Physicians don't follow up when their patients are benefitting from CAM therapies

Physicians may not often pursue information about CAM therapies even when their patients appear to be benefitting from them. If a patient is pursuing a CAM therapy on his/her own and the patient is doing much better than expected, a physician might sometimes acknowledge "a 'between you and me its working' type dialogue" and convey that "'whatever it is, keep doing it'". Even though this type of response acknowledges and legitimizes what the patient is doing in some sense, it does so in a way where there is "still very little dialogue about the paradigmatic basis of the therapy being used" (Tovey and Broom 2007). In cases like this, there is not usually a sense that others, even the physician's other patients, ought to know about the treatment as an option. Presumably this is because of the therapy's unproven, low scientific status. However, since, as discussed, physicians prescribe and support many non-CAM therapies with unproven, low scientific status, science may not be the real issue here.

Physicians may fear for their reputations if they more publicly acknowledge a fringe therapy that may or may not be effective compared to the placebo effect. They may also fear that recommending a fringe therapy to others based on the experience of a single patient would be outside the scope of the "standard of care" and put them at legal risk. So their decision to keep the potential benefit of a CAM therapy between them and the patient could be based on fear and caution for their reputation. At the same time, one would expect a more civic-minded physician to perhaps look into the therapy further and lend some kind of support for further research in it, since it might have real benefits for the physicians' other patients as well. Yet the reality may be that physicians are less inclined to conduct research into an unconventional therapy much further. They might perceive any CAM therapy as outside their field of expertise. And even if they did try to research it, and they limited themselves to highest-standards scientific literature

published in high-quality journals, they would most likely find negative or null/inconclusive placebo-controlled RCTs or meta-analyses, or studies critical of the underlying unscientific rationale of the CAM therapies. Seeing this, they could fall back on the conclusion that the therapy is unproven.

4.4.3 Sociological Issues and Attitudes: Things simply don't get studied even if they are promising

The evidence-based medicine framework only works well if therapies worth studying get studied. When someone says that there is no evidence to support an unconventional therapy, this is often true if the implied meaning of the term "evidence" is highest-standards scientific evidence. The reason why there is often no highest-standards scientific evidence to support unconventional therapies may be due to the fact that no one has ever had the incentive to study them adequately. Adhering to the principles of EBM means that everyone waits for other researchers to pursue promising therapies and present them in the form of published peer-reviewed articles. However, the EBM hierarchy fosters a negative attitude towards observational case reports about unconventional therapies. The general perception towards anecdotes is that they are not data because they involve too much uncertainty; any reported positive effect for a therapy could have been the placebo effect, or regression to the mean, or spontaneous remission, or a misdiagnosis. Therefore, somebody in the world who is scientifically minded would have to go through the effort to get approved and pay for and conduct and write and get peer reviewed and get published a trial of a therapy that controlled for these other factors.

The rational-skeptic attitude of the EBM movement only works well if observationally promising treatments create enough interest by the scientifically-minded where treatments do

eventually get studied in this way. However, the attitude of the evidence-based medicine movement is not conducive to fostering interest in observational case reports. "The proponents of EBM consider case reports and case series as the lowest form of scientific evidence. This opinion has influenced medical editorship and has made the publication of observational reports and studies in first-line journals increasingly difficult" (Berguer 2004). Observational case reports for unconventional therapies face even larger obstacles than the obstacles faced by conventional case studies because of the way unconventional therapies are perceived in the scientific community. "All the rules of academic medicine, from peer review through tenure and promotion criteria, and the operating procedures of NIH, generally guarantee that radically unconventional topics receive little attention and that when studies are done they are rarely published" (Hufford 2003). This effect could just be an internalization of the biological implausibility argument towards CAM therapies, but lack of attention towards unconventional anecdotes may not be compatible with the ultimate goals of the evidence-based medicine movement. If unconventional anecdotes do not generate interest among scientists sociologically, then the minimal requirements for a therapy to get incorporated into the EBM system, even at the level of published case studies, will successfully filter out bad therapies but will also not incorporate potentially beneficial therapies.

Even when potentially beneficial CAM therapies sometimes do get studied because of positive observational reports, the studies might not be conducted properly: the study could have been performed with an inappropriately rigorous method, out of context or on a group where an undetectable subset of participants would experience effective results. If this happens too often, then unconventional medicine practitioners will rightfully become skeptical of the scientific testing process. "Alternative medicine in the past has shown little interest in producing RCT

evidence. Its proponents are less embedded within a science-based epistemology, there is no money available, and there has been an awareness of the limitations of such methodology for studying complex individualised treatments" (Barry 2006). On the other hand, if medical scientists exhibited a more accommodating attitude towards unconventional anecdotes then more therapies worth studying might end up getting studied properly. This improved attitude might go along the lines of, "if the treatment performed in this anecdotal report wasn't caused by the placebo effect or regression to the mean, and if it is safe, then maybe it could be a promising therapy." Unfortunately, this attitude may conflict with the hierarchy of evidence-based medicine as well as the attitude of all scientific testing in general.

Another reason why the evidence-based medicine movement's attitude is problematic despite the EBM hierarchy's theoretical flexibility is that there are different kinds of observational evidence that can serve different purposes. "Between the esoterica of molecular biology and genetics on one side and the recommendations issuing from EBM activity on the other, there is a wide space that should be occupied by clinical intelligence of the proven, and so far most productive, kind: studying the mechanics of disease or the effect of new drugs by prepared observation, inductive hypotheses, and analysis of the value of treatments through the strict study of subsets of patients and physiopathologic reasoning" (Berguer 2004). Many of the greatest scientific discoveries were made without resorting to rigorous randomized testing.

4.4.4 CAM as a Paradigm

In many ways CAM resembles a pre-paradigm within Thomas Kuhn's paradigm framework. Kuhn notes that "[t]he pre-paradigm period, in particular, is regularly marked by frequent and deep debates over legitimate methods, problems, and standards of solution, though these serve rather to define schools than to produce agreement." He cites this happening in debates in optics and electricity as well as the development of seventeenth-century chemistry and early nineteenth-century geology (Kuhn 1962).

Along these lines, Kaptchuk and Miller describe in some detail how CAM has scientifically defined itself in a haphazard manner over time with different methods, problems and solutions:

> Proponents of CAM systems usually assert that they operate within a theoretical and rational understanding of the world validated by the reliability of ordinary human experience. For many alternative medical systems, the process of validation and new discovery involves tension between the teachings of a revered founder (e.g., D.D. Palmer for chiropractic, Samuel Hahnemann for homeopathy) or canonical text (e.g., the Yellow Emperor's Classic of Internal Medicine for East Asian medicine, the Charaka Samhita for India's Ayurveda) and a slow accretion of experience and interpretation, which necessarily selectively emphasizes or neglects aspects of the original sources. This tension between the golden past and the subsequent accumulated experience creates a system that is preserved or rearranged as needed, often under the guidance of authoritative interpreters…Change is slow, uneven, erratic, interpretative, often unacknowledged, and can seem arbitrary, at least to an outsider.
>
> Fear of bias does not haunt CAM. Fundamental to these CAM systems is the belief that simple observation and experience corroborates, elaborates and extends the revered theories and healing practices. "Unimpeachable testimonials" of cures are acceptable evidence; case reports narrated in the singular are acceptable units of authentication. Alternative medicine makes no rigid separation between objective and subjective evidence. Immediate and personal experiences are positively valued, while objective detachment and analytic methods are not. CAM does not rely on evaluations of treatments by experimental controls, blind assessment, placebo comparisons, and statistical inferences, even though such methods are considered essential for adherents of mainstream medicine who want to evaluate CAM therapies.

> Often, vitalistic forces undetectable by natural science that operate according to what is described as "natural law" are postulated to be the basis of observed changes in illness and health (for example, Ayurveda's prana or chiropractic's innate intelligence). Some CAM theories and practices are overtly supernatural. Magical or numinous healing powers are ascribed to objects, people or places. Examples include crystals, amulets, psychic clairvoyants, or healing locations. Some CAM practices ascribe unlimited therapeutic power to the mind (Kaptchuk and Miller 2005).

Some commentators have argued that a CAM research culture should be promoted and that CAM needs to engage the medical community in a reevaluation of its scientific standards.

> "Alternative medical researchers will only ever achieve the most significant goals of integration, respect, and legitimacy within the medical community if they involve themselves in critique and reevaluation of the standards to which they are being held… Recognizing the need for critical discussion and engaging mainstream physicians in a collaborative evaluation of current evidential standards will, I suggest, lead to new and better definitions of good evidence, that are scientifically rigourous and yet responsive to the real needs of practitioners of all types (Borgerson 2006).

> In order to promote the growth of a CAM research culture, the attitudes among CAM practitioners and the skepticism of research ethics committees both need to be overcome. This will involve a process of learning and education. CAM practitioners have to understand the value of science and the fact that without scientific proof CAM is unlikely to survive. Scientifically minded evaluators of CAM projects should understand the nature of CAM and the fact that, in certain instances, scientific rigor can only be taken to a certain point (Ernst, Cohen and Stone 2004).

These two quotations could be considered suggestions that attempt to bridge two paradigms, biomedicine and CAM, together. However, if CAM and conventional medicine are different paradigms in the full Kuhnian sense then they could be what Kuhn calls incommensurable, meaning that it could be impossible for there to be enough of an informational or scientific bridge between the two. On the other hand, if CAM and conventional medicine merely resemble Kuhnian paradigms, there may be some hope for collaboration between the two.

If CAM resembles a pre-paradigm, a more robust CAM paradigm might have a better understanding of the mechanisms and theories behind many CAM therapies. CAM resembles a Kuhnian paradigm because it appears to have different quasi-metaphysical commitments,

different types of data and different puzzles to be solved compared to the current biomedical paradigm. Paul Starr, a medical historian, characterized some of the fundamental assumptions, beliefs and faiths that people tend to have in the current prevailing biomedical paradigm:

> Belief in medicine is the result of many factors:
> years of conditioning by the medical profession;
> a research and philanthropy establishment deeply committed to the medical model;
> a few spectacular medical successes;
> a general, society-wide glorification of science and technology;
> and the medicalization of society, the cultural roots of which extend far beyond professional imperialism (Starr 1976).

Most CAM therapies are based on the existence and manipulation of some kind of vital force. All of CAM's vital force concepts might plausibly be classified as properties of the world and the human body that go beyond current scientific understandings in physics and biology. A new CAM paradigm would probably embrace and incorporate these concepts. It might also have a better and different conception of the placebo effect and the mind-body connection.

4.4.4.1 Studying Qi

The Chinese Medicine vital energy concept of Qi could be translated as "'a subtle global operating system,..human software not human hardware' or as 'a generative matrix in which all things interact with all other things through the exchange of information'" (Shea 2006). Since the 1950s, researchers have been trying to find a tubing structure (like nerve fibers or blood vessels) below the route of the acupuncture meridian systems to find an anatomical correlate or physical basis for the existence of Qi meridians, but such work was in vain (Fan 2003). The future may depend on more time, money and attention on trying to figure out Qi-like phenomenon. This could involve "bench research" at the lower level of the EBM hierarchy.

There has been some empirical research into internal (meridians) and external (qigong healing) Qi phenomena. In 1978 a prominent Chinese physicist at the Shanghai Nuclear

Research Center published some studies of two qigong masters where a special sensor picked up infrared and electromagnetic energy as the masters tried to emit energy from their hands. “The authors of the article concluded that there was nothing going on here contrary to the natural law – in that sense, nothing ‘paranormal’. Qi was a form of energy that was currently unknown, but that eventually would be wholly susceptible to explanation in the language of physics” (Harrington 2001). It is not clear if there was any follow-up to this specific research, but there have not been many well-known scientific studies of Qi phenomenon since then. "[S]cience has not concerned itself with the investigation of Qi –some scientific evaluation of acupuncture effectiveness, yes, but even the mere existence of Qi, no. Much less does science have evidence that contradicts the theory of Qi—that is, firmly supported statements which would be negated by the existence of Qi, its role in health and illness or its manipulation by needling, burning moxa, Qi Gong, etc” (Hufford 2002). There have been some studies on Qi-related healing treatments (Chen 2004), but a strong research base for Qi phenomena is lacking.

4.4.4.2 Incomplete Nature of Current Medical Knowledge

Despite the successes of medical science, the medical community still strives and struggles to understand many chronic human illnesses and diseases. The current biomedical paradigm grapples with many medical conditions. “While [modern medicine] succeeds dramatically in many areas, such as the control of acute and infectious disease, it also fails in others, such as the management of chronic conditions with complex etiologies” (Jonas 2002). Certain sets of symptoms are referred to as syndromes because they are not even conceptually coherent within modern biomedicine. Human mind and brain sciences are not nearly as advanced as one might think they are. “Even today, there are anomalies and internal divisions in the brave new world of brain science that indicate that we are actually not as close to fulfilling the hopes of the first bold

generation of brain researchers as we might sometimes think" (Harrington 2007). Current testing methods still cannot detect many debilitating disorders and there are few comprehensive explanations for many mental illnesses and also few comprehensive explanations for the placebo effect or the relationship between the mind and the body. Recent large-scale attempts to understand the brain are not well-defined and outcomes cannot be easily recognized (Bargmann and Marder 2013). Overall, there appear to be vast amounts of information that have not yet been discovered about the human body and human health.

Medicine became expressly scientific in 1911 after the incorporation of the Flexner Report (Flexner 1910). It has made a lot of progress since the days of bloodletting and frontal lobotomies. The germ theory of disease, antibiotics, genetics and antidepressants are amazing innovations that have been made in the last 150 years. However, the list of mystifying illnesses today seems large enough to at least warrant an examination of the imperfections of current information gathering practices in medicine and an entertainment of alternative attitudes or alternative methods of inquiry with regard to scientific testing. It is not clear if scientific medicine has successfully incorporated most of the useful information about medical disorders that has been generated throughout history.

4.4.4.3 What a CAM paradigm would look like

A treatment in a new paradigm that embraced some CAM concepts that are beyond current science might not need to prove that the therapy really "does anything" compared to a placebo control group. It would be assumed that it does do something, at least for some people, and priority would be given to finding ways to maximize healing with it. More priority might be given to pragmatic research, subjective reports of health and illness, trial-and-error and observational reports. A CAM paradigm would have different fundamental assumptions about

the human body and the world, and perhaps a greater understanding of the brain, the mechanism of the placebo effect and the mechanism of the mind-body connection. Medical therapies would address and target the underlying causes of many currently unsolvable or untreatable illnesses and be able to reduce the severity of them more successfully than what conventional medicine has been able to do for them.

4.4.5 Money and Politics

One criticism of applying the framework of EBM to CAM has to do with the politics of biomedicine:

> "Mike Saks suggests RCT evidence is being used strategically by biomedicine's medical associations, to reduce potential threat from alternative medicines, feeding into incorporationist and assimilationist policies (Saks, 1996). In one case study Saks charts how the British Medical Association (BMA) moved from a completely hostile position of total exclusion of alternative medicine (in their 1986 report) to an idea of embracing collaboration with non-medically qualified alternative practitioners (in their 1993 report). Saks interprets this U-turn as a heavily defensive, politically motivated strategy, aimed at reducing the potential threat to biomedicine's position, in terms of power, income and status, from alternative therapies. Baer (2004) has noted similar shifts within the American Medical Association, and interpreted them similarly" (Barry 2006).

This political-economic line of argument may have some amount of validity to it because perceptions did apparently change from "not studied and not worth studying" to "not studied and worth studying", and it might be possible to interpret this change in perception as an attempt to address a growing threat that can no longer be ignored. Additionally, many physicians can now point to the National Center for Complementary and Alternative Medicine (NCCAM) within the National Institutes of Health to deflect arguments from CAM proponents that CAM therapies do not receive enough research attention. However, NCCAM has studied a small number of

untraditional therapies and the way that CAM therapies have been studied in the past may not be the optimal way to study them.

On the other hand, the manner in which CAM therapies have become popular in the United States does not lend much support to political arguments about biomedical attempts to control and suppress CAM. The increasing use of alternative medicine in the United States (Tindle et al 2005), the fact that the Office of Alternative Medicine (now the NCCAM) in the National Institutes of Health was established at the behest of a small number of enthusiastic senators (Young 1998) and the fact that many CAM centers inside hospitals were established by wealthy private benefactors (Atwood 2004) indicates that smaller interests may have succeeded in bringing CAM into the public consciousness to the point where it became apparent to some that the therapies really were worth studying. Additionally, it is unsurprising that authorities within biomedicine would call for non-biomedical therapies to meet the same standards as conventional therapies. If biomedicine is considered a Kuhnian paradigm, it should be expected that "normal scientists" inside a paradigm would assume that their own standard methods of evaluation would apply to all therapies (Kuhn 1962).

4.4.6 Funding

Even though some wealthy benefactors have funded Integrative Medicine centers in hospitals and the NIH gives some money to the National Center for Complementary and Alternative Medicine, an oft-cited argument is that there is not much of a profit motive to study CAM therapies. Clinical studies can take an enormous amount of effort and money to conduct. The FDA drug approval process takes six to fifteen years and costs $100 million to $880 million per drug (Johns 2006). No matter how much money is offered to study CAM therapies through

PCORI or elsewhere, it may never be enough to thoroughly evaluate most of it. “It will only ever be practical to subject a limited number of items to experimental evaluation. We therefore need to take advantage of other methods to try and fill in the huge gaps that are always likely to exist in the experimental published findings” (Black 1996). It is not realistic to expect that anyone will eventually get around to properly studying most CAM therapies anytime soon.

5.0 CONCLUSION

There is no established method of studying CAM, and current methods may suffer from paradigm-related problems in conducting clinical trials and interpreting evidence. It is not clear if the attitude of the EBM hierarchy that favors RCTs offers much hope of determining how to study CAM. Its priorities are heavily weighted towards rigor and efficacy but, assuming CAM therapies have real effects, the way forward may be weighted towards pragmatic outcomes research, or even the middle and bottom of the EBM hierarchy with observational trials and case studies. The priorities of the EBM movement may need to be altered to properly accommodate the epistemological issues raised by CAM therapies.

Many health professionals who analyze how to research complementary and alternative medicine conclude that the "same" methods of conventional medicine can and should be used for unconventional medicine. However, these comments often do not specify what conventional research methods entail. "[T]alking about evidence without being sure what methods, techniques, and strategies can be relied upon to produce valid evidence is talking through your epistemological hat…few physicians or those who work with them have any sophistication about the philosophy of science. Even fewer have ever been taught anything about the philosophy of medicine. And fewer still can give a coherent presentation on what the core infrastructure is that distinguishes the science of medicine from the faith and testimonials of religious healers or the loopy claims of the talk radio nutritionists." (Caplan 2006). This quotation from Bioethicist

Arthur Caplan is critical of people who talk about evidence without knowing enough about the philosophy of science and the philosophy of medicine, but even Prof. Caplan's quotation seems to be presupposing that the science of medicine has a strong, solid foundation. Yet the imperfections of the evidence-based medicine framework, including those illuminated by the challenges of trying to rigorously study complementary and alternative medicine, may indicate that researchers will always be talking through their epistemological hats. "Studies of CAM bring the whole biomedical research apparatus into question, particularly because, while there is clearly knowledge to be gained from studying CAM, there is no established method of proceeding" (Derkatch 2009). It may be the case that no one can really ever be sure what methods can be relied upon to produce valid evidence. Can pragmatic "whole systems" comparative effectiveness outcomes research validate therapies like traditional Chinese medicine that incorporate potentially unfalsifiable quasi-metaphysical commitments? Can dramatic anecdotes validate unconventional therapies that have not been properly studied and are not going to be properly studied in the near future? The hierarchy of evidence-based medicine does not have good answers to these questions and the attitude of the evidence-based medicine movement offers little hope of answering them. Self-critical conceptual shifts in medical research may be needed to accommodate the problematic priorities of the EBM hierarchy and research obstacles posed by complementary and alternative medicine therapies.

5.1 RESEARCH AND POLICY RECOMMENDATIONS

If modern biomedical scientific methods are going to be used to try to verify and validate complementary and alternative medicine practices, the research and testing methods need to be

more flexible and creative. One example of a creative study is a 2012 meta-analysis that analyzed raw trial data, rather than published results, from 29 sham/placebo-controlled acupuncture RCTs. The study found a "modest" but statistically significant difference between real acupuncture and sham/placebo acupuncture based on the individual data from about 18,000 trial subjects (Vickers et al 2012).

The following are some recommendations for the future of CAM research and medical research in general:

5.1.1 For CAM research

1. Pragmatic comparative effectiveness outcomes research should be conducted on CAM therapies more often than placebo-controlled efficacy trials. Patterns of dramatic effects that are not normally achieved with conventional treatments should be pursued and highlighted.

2. Physicians should be less hesitant and more open to recommend CAM therapies that have positive evidence from outcomes-focused comparative effectiveness studies, positive effects from N of 1 trials or patterns of dramatic effects, as long as there is some confidence in the CAM therapies' safety.

3. There should be more research exploring the underlying mechanisms of CAM therapies and CAM phenomena such as Qi, acupuncture meridians, human biofields, chakras, and water memory. A plurality of methods should be employed in addition to clinical trials: correlations with changes in objective biological markers, other bench research,

ethnographies, investigative commissions of experts, qualitative and quantitative research, observational studies, case reports, case series studies, etc.

4. There should also be more research exploring the mechanism of the placebo effect, the mind-body connection and other non-specific effects.

5. Placebo-controlled efficacy trials for CAM therapies need to incorporate better placebo controls, account for the belief prerequisite, and account for the possibility of unidentified subsets. Until this can be achieved and until debates over efficacy are more settled, pragmatic effectiveness outcomes studies as well as N of 1 studies and studies with dramatic effects should be considered the "current best evidence in making decisions about the care of individual patients" for CAM therapies.

6. Case studies and case series studies of successful healings possibly attributable to CAM therapies should be encouraged and published more often. This would identify therapies to pursue further with observational studies or effectiveness outcomes trials.

5.1.2 For medical research in general

7. Research methods that don't use placebo controls and randomization can still be considered good research. Sometimes the results from a variety of research methods that are lower on the evidence-based medicine hierarchy (pragmatic trials, observational studies, evidence of biological effects, case studies) can be more informative than the results from trials that have placebo controls and randomization.

8. If a rigorous randomized placebo-controlled clinical trial is negative or null/inconclusive, it should not cause a therapy to be considered definitively “disproven” or “proven ineffective” if there are other types of evidence in support of the treatment.

BIBLIOGRAPHY

AcademyHealth. (2009). *A First Look at the Volume and Cost of Comparative Effectiveness Research in the United States*. Retrieved from http://www.academyhealth.org/files/FileDownloads/AH_Monograph_09FINAL7.pdf

Angell, M. (2011, June 23). The Epidemic of Mental Illness: Why? *The New York Review of Books*.

Angell, M., & Kassirer, J. P. (1998). Alternative medicine--the risks of untested and unregulated remedies. *The New England Journal of Medicine*, *339*(12), 839–41. doi:10.1056/NEJM199809173391210

Applications for FDA Approval to Market a New Drug - 21CFR314.126 (2014). Retrieved from http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?fr=314.126

Ashcroft, R. E. (2004). Current epistemological problems in evidence based medicine. *Journal of Medical Ethics*, *30*(2), 131–135. doi:10.1136/jme.2003.007039

Atwell, B. L. (2003). Mainstreaming Complementary and Alternative Medicine in the Face of Uncertainty. *UMKC Law Review*, *72*. Retrieved from http://heinonline.org/HOL/Page?handle=hein.journals/umkc72&id=605&div=&collection=journals

Atwood, K. C. (2004). Naturopathy, pseudoscience, and medicine: myths and fallacies vs truth. *MedGenMed : Medscape General Medicine*, *6*(1), 33. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1140750&tool=pmcentrez&rendertype=abstract

Bargmann, C. I., & Marder, E. (2013). From the connectome to brain function. *Nature Methods*, *10*(6), 483–490. doi:10.1038/nmeth.2451

Barnes, P. M., Bloom, B., & Nahin, R. L. (2008). Complementary and alternative medicine use among adults and children: United States, 2007. *National Health Statistics Reports*, (12), 1–23. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/19361005

Barry, C. A. (2006). The role of evidence in alternative medicine: contrasting biomedical and anthropological approaches. *Social Science & Medicine*, *62*(11), 2646–57. doi:10.1016/j.socscimed.2005.11.025

Berguer, R. (2004). The evidence thing. *Annals of Vascular Surgery*, *18*(3), 265–70. doi:10.1007/s10016-004-0044-3

Bishop, T. F., Federman, A. D., & Keyhani, S. (2010). Physicians' views on defensive medicine: a national survey. *Archives of Internal Medicine*, *170*(12), 1081–3. doi:10.1001/archinternmed.2010.155

Black, N. (1996). Why we need observational studies to evaluate the effectiveness of health care. *BMJ*, *312*(7040), 1215–1218. doi:10.1136/bmj.312.7040.1215

Bluhm, R. (2005). From hierarchy to network: a richer view of evidence for evidence-based medicine. *Perspectives in Biology and Medicine*, *48*(4), 535–47. doi:10.1353/pbm.2005.0082

Bluhm, R. (2009). Some observations on observational research. *Perspectives in Biology and Medicine*, *52*(2), 252–63. doi:10.1353/pbm.0.0076

Bone, K., & Mills, S. (2013). *Principles and Practice of Phytotherapy: Modern Herbal Medicine*. Elsevier Health Sciences. Retrieved from http://books.google.com/books?id=Ca3QAQAAQBAJ&pgis=1

Borgerson, K. (2005). Evidence-based alternative medicine? *Perspectives in Biology and Medicine*, *48*(4), 502–15. doi:10.1353/pbm.2005.0084

Borgerson, K. (2006). Playing by the Rules: Feminism, Alternative Medicine and Standards of Evidence. In J. C. Cohen & J. E. Keelan (Eds.), *Comparative Program on Health and Society Lupina Foundation Working Papers Series 2004–2005*. The Comparative Program on Health and Society.

Borgerson, K. (2009). Valuing evidence: bias and the evidence hierarchy of evidence-based medicine. *Perspectives in Biology and Medicine*, *52*(2), 218–33. doi:10.1353/pbm.0.0086

Bouchet, C., Guillemin, F., & Briançon, S. (1996). Nonspecific effects in longitudinal studies: Impact on quality of life measures. *Journal of Clinical Epidemiology*, *49*(1), 15–20. doi:10.1016/0895-4356(95)00540-4

Brandt, A. M. (1990). The Cigarette, Risk, and American Culture. *Daedalus*, *119*(4), 155–176.

Briggs, J. P. (2014, January 2). When To Get Pragmatic. Retrieved from http://nccam.nih.gov/research/blog/pragmatictrials

Brody, H. (2002). The placebo effect: implications for the study and practice of complementary and alternative medicine. In D. Callahan (Ed.), *The Role of Complementary and Alternative Medicine: Accommodating Pluralism*. Washington, DC: Georgetown University Press.

Caplan, A. L. (2006). No Method, Thus Madness? *Hastings Center Report*, *36*(2), 12–13. doi:10.1353/hcr.2006.0019

Chen, K. W. (2004). An analytic review of studies on measuring effects of external QI in China. *Alternative Therapies in Health and Medicine*, *10*(4), 38–50. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15285273

Cherkin, D. C., Sherman, K. J., Avins, A. L., Erro, J. H., Ichikawa, L., Barlow, W. E., … Deyo, R. A. (2009). A randomized trial comparing acupuncture, simulated acupuncture, and usual care for chronic low back pain. *Archives of Internal Medicine*, *169*(9), 858–66. doi:10.1001/archinternmed.2009.65

Cohen, M. H. (2004). Negotiating Integrative Medicine: A Framework for Provider-Patient Conversations. *Negotiation Journal*, *20*(3), 409–433. doi:10.1111/j.1571-9979.2004.00035.x

Comparative Effectiveness Research. (2014). Retrieved November 13, 2014, from http://tracs.unc.edu/index.php/cer-strategic-initiative

Concato, J., Shah, N., & Horwitz, R. I. (2000). Randomized, controlled trials, observational studies, and the hierarchy of research designs. *The New England Journal of Medicine*, *342*(25), 1887–92. doi:10.1056/NEJM200006223422507

Connor, S. (2003, December 8). Glaxo chief: Our drugs do not work on most patients. *The Independent*.

Derkatch, C. J. (2009, October). *Rhetorical boundaries in the "new science" of alternative medicine*. University of British Columbia. Retrieved from http://circle.ubc.ca/handle/2429/14201

Diener, H.-C., Kronfeld, K., Boewing, G., Lungenhausen, M., Maier, C., Molsberger, A., … Meinert, R. (2006). Efficacy of acupuncture for the prophylaxis of migraine: a multicentre randomised controlled clinical trial. *The Lancet. Neurology*, *5*(4), 310–6. doi:10.1016/S1474-4422(06)70382-9

Ernst, E. (1998). Single-case studies in complementary/alternative medicine research. *Complementary Therapies in Medicine*, *6*(2), 75–78. doi:10.1016/S0965-2299(98)80079-4

Ernst, E. (2002). What's the point of rigorous research on complementary/alternative medicine? *Journal of the Royal Society of Medicine*, *95*(4), 211–3. Retrieved from http://jrs.sagepub.com/content/95/4/211.full

Ernst, E., Cohen, M. H., & Stone, J. (2004). Ethical problems arising in evidence based complementary and alternative medicine. *Journal of Medical Ethics*, *30*(2), 156–159. doi:10.1136/jme.2003.007021

Ernst, E., Lee, M. S., & Choi, T.-Y. (2011). Acupuncture: does it alleviate pain and are there serious risks? A review of reviews. *Pain*, *152*(4), 755–64. doi:10.1016/j.pain.2010.11.004

Fan, R. (2003). Modern Western Science as a Standard for Traditional Chinese Medicine: A Critical Appraisal. *The Journal of Law, Medicine & Ethics*, *31*(2), 213–221. doi:10.1111/j.1748-720X.2003.tb00082.x

FINAL REPORT of the White House Commission on Complementary and Alternative Medicine Policy. (2002). Retrieved from http://www.whccamp.hhs.gov/es.html

Flexner, A. (1910). *Medical education in the United States and Canada: a report to the Carnegie Foundation for the Advancement of Teaching*. Carnegie Foundation for the Advancement of Teaching. Retrieved from http://books.google.com/books?hl=en&lr=&id=FDZrAAAAMAAJ&pgis=1

Fontanarosa, P. B., & Lundberg, G. D. (1998). Alternative Medicine Meets Science. *JAMA*, *280*(18), 1618. doi:10.1001/jama.280.18.1618

Georgiou, H. (2006). *Reasons for use and disclosure of complementary medicine by people with haemoglobinopathy*. Retrieved from http://vuir.vu.edu.au/1479/1/Georgiou.pdf

Glionna, J. M. (1996, September 17). Turning East for the Answers to Medical Mysteries. *The Los Angeles Times*.

Gøtzsche, P. C. (2012). Why we need to open up health research by sharing our raw data. In *Opening up medical research data for an ethical and efficient EU policy*. Copenhagen, Denmark.

Greenhalgh, T. (2010). *How to Read a Paper: The Basics of Evidence-Based Medicine* (Vol. 18). John Wiley & Sons. Retrieved from http://books.google.com/books?id=q-92iFMFJGwC&pgis=1

Grossman, J., & Mackenzie, F. J. (2005). The randomized controlled trial: gold standard, or merely standard? *Perspectives in Biology and Medicine*, *48*(4), 516–34. doi:10.1353/pbm.2005.0092

Guyatt, G. et al. (1992). Evidence-based medicine: a new approach to teaching the practice of medicine. *JAMA*, *268*(17), 2420. doi:10.1001/jama.1992.03490170092032

Guyatt, G., Rennie, D., Meade, M., & Cook, D. (2008). *Users' Guides to the Medical Literature: Essentials of Evidence-Based Clinical Practice, Second Edition: A Manual for Evidence-Based Clinical Practice, Second Edition Ebook*. McGraw Hill Professional. Retrieved from

http://books.google.com/books/about/Users_Guides_to_the_Medical_Literature_E.html?id=Kz3_36ejIv4C&pgis=1

Haake, M., Müller, H.-H., Schade-Brittinger, C., Basler, H. D., Schäfer, H., Maier, C., … Molsberger, A. (2007). German Acupuncture Trials (GERAC) for chronic low back pain: randomized, multicenter, blinded, parallel-group trial with 3 groups. *Archives of Internal Medicine*, *167*(17), 1892–8. doi:10.1001/archinte.167.17.1892

Hammerschlag, R., & Zwickey, H. (2006). Evidence-based complementary and alternative medicine: back to basics. *Journal of Alternative and Complementary Medicine*, *12*(4), 349–50. doi:10.1089/acm.2006.12.349

Harrington, A. (2001). China, "Qi", and the challenges of engaged scholarship. In G. Allen & R. M. MacLeod (Eds.), *Science, History and Social Activism*. Springer Science & Business Media.

Harrington, A. (2007). How to house a mind inside a brain. Lessons from history. *EMBO Reports*, *8 Spec No*(1S), S12–5. doi:10.1038/sj.embor.7400974

Haynes, R. B. (2002). What kind of evidence is it that Evidence-Based Medicine advocates want health care providers and consumers to pay attention to? *BMC Health Services Research*, *2*(1), 3. doi:10.1186/1472-6963-2-3

Hempel, C. G. (1966). Philosophy of Natural Science. Retrieved from http://philpapers.org/rec/HEMPON

Henry, S. G., Zaner, R. M., & Dittus, R. S. (2007). Viewpoint: Moving beyond evidence-based medicine. *Academic Medicine : Journal of the Association of American Medical Colleges*, *82*(3), 292–7. doi:10.1097/ACM.0b013e3180307f6d

Horn, S. D., DeJong, G., Ryser, D. K., Veazie, P. J., & Teraoka, J. (2005). Another look at observational studies in rehabilitation research: going beyond the holy grail of the randomized controlled trial. *Archives of Physical Medicine and Rehabilitation*, *86*(12 Suppl 2), S8–S15. doi:10.1016/j.apmr.2005.08.116

Howick, J., Chalmers, I., Glasziou, P., Greenhalgh, T., Heneghan, C., Liberati, A., … Thornton. Hazel. (2011). *Explanation of the 2011 Oxford Centre for Evidence-Based Medicine (OCEBM) Levels of Evidence (Background Document)*. Retrieved from http://www.cebm.net/ocebm-levels-of-evidence/

Howick, J. H. (2011). *The Philosophy of Evidence-based Medicine*. John Wiley & Sons. Retrieved from http://books.google.com/books?hl=en&lr=&id=O8djbHBva5IC&pgis=1

Hrobjartsson, A., & Brorson, S. (2002). Interpreting results from randomized clinical trials of complementary/alternative interventions: The role of trial quality and pre-trial beliefs. In D.

Callahan (Ed.), *The Role of Complementary and Alternative Medicine: Accommodating Pluralism*. Washington, DC: Georgetown University Press.

Hufford, D. (2002). CAM and Cultural Diversity: Ethics and Epistemology Converge. In D. Callahan (Ed.), *The Role of Complementary and Alternative Medicine: Accommodating Pluralism*. Washington, DC: Georgetown University Press.

Hufford, D. (2003). Evaluating Complementary and Alternative Medicine: The Limits of Science and of Scientists. *The Journal of Law, Medicine & Ethics*, *31*(2), 198–212. doi:10.1111/j.1748-720X.2003.tb00081.x

Hughes, B. M. (2008). How should clinical psychologists approach complementary and alternative medicine? Empirical, epistemological, and ethical considerations. *Clinical Psychology Review*, *28*(4), 657–75. doi:10.1016/j.cpr.2007.09.005

Imrie, R., & Ramey, D. W. (2000). The evidence for evidence-based medicine. *Complementary Therapies in Medicine*, *8*(2), 123–6. doi:10.1054/ctim.2000.0370

Ioannidis, J. P. A. (2005). Contradicted and initially stronger effects in highly cited clinical research. *JAMA*, *294*(2), 218–28. doi:10.1001/jama.294.2.218

Jensen, U. J. (2007). The Struggle for Clinical Authority: Shifting Ontologies and the Politics of Evidence. *BioSocieties*, *2*(1), 101–114. doi:10.1017/S174585520700508X

Johns, M. Z. (2006). Informed Consent: Requiring Doctors to Disclose Off-Label Prescriptions and Conflicts of Interest. *Hastings Law Journal*, *58*. Retrieved from http://heinonline.org/HOL/Page?handle=hein.journals/hastlj58&id=999&div=&collection=journals

Jonas, W. (2002). Evidence, Ethics and the Evaluation of Global Medicine. In D. Callahan (Ed.), *The Role of Complementary and Alternative Medicine: Accommodating Pluralism*. Washington, DC: Georgetown University Press.

Kaptchuk, T. J., Kelley, J. M., Conboy, L. A., Davis, R. B., Kerr, C. E., Jacobson, E. E., … Lembo, A. J. (2008). Components of placebo effect: randomised controlled trial in patients with irritable bowel syndrome. *BMJ (Clinical Research Ed.)*, *336*(7651), 999–1003. doi:10.1136/bmj.39524.439618.25

Kaptchuk, T. J., & Miller, F. G. (2005). Viewpoint: What is the Best and Most Ethical Model for the Relationship Between Mainstream and Alternative Medicine: Opposition, Integration, or Pluralism? *Academic Medicine*, *80*(3), 286–290. doi:10.1097/00001888-200503000-00015

Kendall, T., & Cochrane, A. (2007). Truth, Disclosure and the Influence of Industry on the Development of NICE Guidelines: An Interview with Tim Kendall. *BioSocieties*, *2*(1), 129–140. doi:10.1017/S1745855207005108

Kolata, G., & Pollack, A. (2008, July 6). Costly Cancer Drug Offers Hope, but Also a Dilemma. *The New York Times*.

Kuhn, T. S. (1962). *The Structure of Scientific Revolutions*. University of Chicago Press. Retrieved from http://books.google.com/books?id=xnjS401VuFMC&pgis=1

Langevin, H. M., Wayne, P. M., Macpherson, H., Schnyer, R., Milley, R. M., Napadow, V., … Hammerschlag, R. (2011). Paradoxes in acupuncture research: strategies for moving forward. *Evidence-Based Complementary and Alternative Medicine : eCAM*, *2011*, 180805. doi:10.1155/2011/180805

Lehrer, J. (2010, December 13). The Truth Wears Off: Is There Something Wrong With the Scientific Method? *The New Yorker*.

Lewith, G. T., Jonas, W. B., & Walach, H. (2010). *Clinical Research in Complementary Therapies: Principles, Problems and Solutions* (Vol. 23, p. 426). Elsevier Health Sciences. Retrieved from http://books.google.com/books?hl=en&lr=&id=CSNw-spnFdkC&pgis=1

Lieberthal, A. S., Carroll, A. E., Chonmaitree, T., Ganiats, T. G., Hoberman, A., Jackson, M. A., … Tunkel, D. E. (2013). The diagnosis and management of acute otitis media. *Pediatrics*, *131*(3), e964–99. doi:10.1542/peds.2012-3488

Lohr, J. M., Demaio, C., & Mcglynn, F. D. (2003). Specific and Nonspecific Treatment Factors in the Experimental Analysis of Behavioral Treatment Efficacy. *Behavior Modification*, *27*(3), 322–368. doi:10.1177/0145445503253831

MacPherson, H., Peters, D., & Zollman, C. (2009). Closing the evidence gap in integrative medicine. *BMJ (Clinical Research Ed.)*, *339*(sep01_2), b3335. doi:10.1136/bmj.b3335

Meisel, Z. F., & Karlawish, J. (2011). Narrative vs evidence-based medicine--and, not or. *JAMA*, *306*(18), 2022–3. doi:10.1001/jama.2011.1648

Morreim, E. H. (2003). A Dose of Our Own Medicine: Alternative Medicine, Conventional Medicine, and the Standards of Science. *The Journal of Law, Medicine & Ethics*, *31*(2), 222–235. doi:10.1111/j.1748-720X.2003.tb00083.x

Nahin, R. L., Barnes, P. M., Stussman, B. J., & Bloom, B. (2009). Costs of complementary and alternative medicine (CAM) and frequency of visits to CAM practitioners: United States, 2007. *National Health Statistics Reports*, (18), 1–14. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/19771719

O'Connor, B. B. (2002). Personal Experience, Popular Epistemology, and Complementary/Alternative Medicine Research. In D. Callahan (Ed.), *The Role of Complementary and Alternative Medicine: Accommodating Pluralism*. Washington, DC: Georgetown University Press.

Paterson, C., & Dieppe, P. (2005). Characteristic and incidental (placebo) effects in complex interventions such as acupuncture. *BMJ (Clinical Research Ed.)*, *330*(7501), 1202–5. doi:10.1136/bmj.330.7501.1202

Patient-Centered Outcomes Research Definition Revision: Response to Public Input. (2012). Washington, DC. Retrieved from http://www.pcori.org/assets/PCOR-Definition-Revised-Draft-and-Responses-to-Input.pdf

Petryna, A. (2007). Clinical Trials Offshored: On Private Sector Science and Public Health. *BioSocieties*, *2*(1), 21–40. doi:10.1017/S1745855207005030

Pollock, A. (2008, December 29). Patient's DNA May Be Signal to Tailor Medication. *The New York Times*.

PROspective Multicenter Imaging Study for Evaluation of Chest Pain. (n.d.). Retrieved November 13, 2014, from http://clinicaltrials.gov/show/NCT01174550

Ratner, R., Eden, J., Wolman, D., Greenfield, S., & Sox, H. (2009). *Initial National Priorities for Comparative Effectiveness Research*. Washington, DC: National Academies Press.

Riveros, C., Dechartres, A., Perrodeau, E., Haneef, R., Boutron, I., & Ravaud, P. (2013). Timing and completeness of trial results posted at ClinicalTrials.gov and published in journals. *PLoS Medicine*, *10*(12), e1001566; discussion e1001566. doi:10.1371/journal.pmed.1001566

Sackett, D. L. (2000). *Evidence-based medicine: how to practice and teach EBM, Volume 2*. Churchill Livingstone. Retrieved from http://books.google.com/books/about/Evidence_based_medicine.html?id=oYJrAAAAMAAJ&pgis=1

Sackett, D. L., Rosenberg, W. M. C., Gray, J. A. M., Haynes, R. B., & Richardson, W. S. (1996). Evidence based medicine: what it is and what it isn't. *BMJ*, *312*(7023), 71–72. doi:10.1136/bmj.312.7023.71

Schaffner, K. F. (2002). Assessments of Efficacy in Biomedicine: the Turn Toward Methodological Pluralism. In Callahan D (Ed.), *The Role of Complementary and Alternative Medicine: Accommodating Pluralism*. Washington, DC: Georgetown University Press.

Scharf, H.-P. (2006). Acupuncture and Knee Osteoarthritis. *Annals of Internal Medicine*, *145*(1), 12. doi:10.7326/0003-4819-145-1-200607040-00005

Schneiderman, L. J. (2003). The (Alternative) Medicalization of Life. *The Journal of Law, Medicine & Ethics*, *31*(2), 191–197. doi:10.1111/j.1748-720X.2003.tb00080.x

Shea, J. L. (2006). Applying evidence-based medicine to traditional chinese medicine: debate and strategy. *Journal of Alternative and Complementary Medicine*, *12*(3), 255–63. doi:10.1089/acm.2006.12.255

Snyder, L. (2007). *Complementary and Alternative Medicine: Ethics, the Patient, and the Physician*. Springer Science & Business Media. Retrieved from http://books.google.com/books?hl=en&lr=&id=spTGldo_XmIC&pgis=1

Solomon, M. (2011). Just a paradigm: evidence-based medicine in epistemological context. *European Journal for Philosophy of Science*, *1*(3), 451–466. doi:10.1007/s13194-011-0034-6

Spear, B. B., Heath-Chiozzi, M., & Huff, J. (2001). Clinical application of pharmacogenetics. *Trends in Molecular Medicine*, *7*(5), 201–204. doi:10.1016/S1471-4914(01)01986-4

Starr, P. (1976). The Politics of Therapeutic Nihilism. *The Hastings Center Report*, *6*(5), 24. doi:10.2307/3561255

Three Approaches to Glucose Monitoring in Non-insulin Treated Diabetes. (n.d.). Retrieved November 13, 2014, from http://clinicaltrials.gov/show/NCT02033499

Timmermans, S., & Berg, M. (2003). *The Gold Standard: The Challenge of Evidence-based Medicine and Standardization in Health Care*. Temple University Press.

Tindle, H. A., Davis, R. B., Phillips, R. S., & Eisenberg, D. M. (2005). Trends in use of complementary and alternative medicine by US adults: 1997-2002. *Alternative Therapies in Health and Medicine*, *11*(1), 42–9. Retrieved from http://europepmc.org/abstract/med/15712765

Tonelli, M. R., & Callahan, T. C. (2001). Why alternative medicine cannot be evidence-based. *Academic Medicine : Journal of the Association of American Medical Colleges*, *76*(12), 1213–20. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11739043

Tovey, P., & Broom, A. (2007). Oncologists' and specialist cancer nurses' approaches to complementary and alternative medicine and their impact on patient action. *Social Science & Medicine*, *64*(12), 2550–64. doi:10.1016/j.socscimed.2007.02.010

Upshur, R. E. G. (2003). Are all evidence-based practices alike? Problems in the ranking of evidence. *Can. Med. Assoc. J.*, *169*(7), 672–673. Retrieved from http://www.cmaj.ca/content/169/7/672.short

Vickers, A. J., Cronin, A. M., Maschino, A. C., Lewith, G., MacPherson, H., Foster, N. E., … Linde, K. (2012). Acupuncture for chronic pain: individual patient data meta-analysis. *Archives of Internal Medicine*, *172*(19), 1444–53. doi:10.1001/archinternmed.2012.3654

Walach, H. (2003). Reinventing the wheel will not make it rounder: controlled trials of homeopathy reconsidered. *Journal of Alternative and Complementary Medicine*, *9*(1), 7–13. doi:10.1089/107555303321222865

Walach, H. (2011). Placebo controls: historical, methodological and general aspects. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *366*(1572), 1870–8. doi:10.1098/rstb.2010.0401

Walach, H., Falkenberg, T., Fønnebø, V., Lewith, G., & Jonas, W. B. (2006). Circular instead of hierarchical: methodological principles for the evaluation of complex interventions. *BMC Medical Research Methodology*, *6*, 29. doi:10.1186/1471-2288-6-29

Wartolowska, K., Judge, A., Hopewell, S., Collins, G. S., Dean, B. J. F., Rombach, I., … Carr, A. J. (2014). Use of placebo controls in the evaluation of surgery: systematic review. *BMJ (Clinical Research Ed.)*, *348*, g3253. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4029190&tool=pmcentrez&rendertype=abstract

Wendland, C. L. (2007). The Vanishing Mother: Cesarean Section and Evidence-Based Obstetrics. *Medical Anthropology Quarterly*, *21*(2), 218–233. doi:10.1525/maq.2007.21.2.218

Wessely, S. (2007). A Defence of the Randomized Controlled Trial in Mental Health. *BioSocieties*, *2*(1), 115–127. doi:10.1017/S1745855207005091

West, S., King, V., Carey, T. S., Lohr, K. N., McKoy, N., Sutton, S. F., & Lux, L. (2002). Systems to rate the strength of scientific evidence. *Evidence Report/technology Assessment (Summary)*, (47), 1–11. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11979732

Whitmarsh T. (2002). The Nature of Evidence in Complementary and Alternative Medicine: Ideas from Trials of Homeopathy in Chronic Headache. In Callahan D (Ed.), *The Role of Complementary and Alternative Medicine: Accommodating Pluralism*. Washington, DC: Georgetown University Press.

Wolford, B. (2014). Big Pharma Plays Hide-the-Ball With Data. *Newsweek*.

Worrall, J. (2007). Evidence in Medicine and Evidence-Based Medicine. *Philosophy Compass*, *2*(6), 981–1022. doi:10.1111/j.1747-9991.2007.00106.x

Young, J. H. (1998). The Development of the Office of Alternative Medicine in the National Institutes of Health, 1991-1996. *Bulletin of the History of Medicine*, *72*(2), 279–298. doi:10.1353/bhm.1998.0110

Ziliak, S. T., & McCloskey, D. N. (2008). *The Cult of Statistical Significance: How the Standard Error Costs Us Jobs, Justice, and Lives*. University of Michigan Press. Retrieved from http://books.google.com/books?hl=en&lr=&id=JWLIRr_ROgAC&pgis=1